Eötvös Loránd Tudományegyetem
Informatikai Kar

Cyclic-Waiting and Vacational Queuing Systems

PhD dissertation

Kárász Péter
ELTE IK Doktori Iskola
(vezető: Dr. Demetrovics János)
Numerikus és szimbolikus számítások doktori program
(programvezető: Dr. Járai Antal)

Témavezető: Dr. Lakatos László
a matematikai tudományok kandidátusa
egyetemi docens

2008

Contents

Chapter 1

Introduction

The Oxford English Dictionary shows the earliest known use of the expression of *queuing system* by David G. Kendall in 1951 in [32], and a use of *queuing theory* in 1954 in Science News. However, the investigation of queues began much earlier than that, the first major study of congestion problems was undertaken by Agner K. Erlang in 1908 under the auspices of the Copenhagen Telephone Company, where he began applying probability to various problems arising in the context of telephone calls. He investigated systems where entering customers form Poisson-processes (the time elapsed between two arrivals is an exponentially distributed random variable), and the service time of customers is also exponentially distributed. He published his first paper on queues in 1909. In 1917 he gave a formula for loss and waiting time which was soon used by telephone companies in many countries. Later it turned out that models elaborated for this purpose can answer questions emerging at other areas of science and technology. In the 1930's Khinchin examined problems within this circle, and the area started to develop steadily from the 50's, when the list of possible applications gradually extended; and which now seems endless, especially since the information technology boom.

Whenever a system is observed where customers come to get a specific service — i. e. they queue up, get serviced and leave the system — and their appearance and service times can be described with probabilistic tools, we talk about a *queuing* (or in other words a *mass-service*) *system*.

It comes naturally to set the question what happens to requests refused because of a busy server. These customers leave the system and return after some time. Possible applications include the landing of aeroplanes or customers at a supermarket who leave non-accessible cash-desks in the hope of later finding a shorter queue. The search for the answer led to the development of a special class of queuing systems, the so-called *retrial systems*. Retrial queues are characterized by the following feature: a customer arriving when all servers accessible for them are busy leaves the service area, but after some random period repeats the request. Thus, the flow of requests in this type of system consists of two parts: the flow of primary requests, which reflects the real wishes of customers, and the flow of repeated requests, which is the consequence of the lack of success of previous attempts. This feature plays a special role in computer and communication networks, but possible applications of such systems are not restricted to them. Standard queuing models do not take into account this repetitive characteristic, and therefore cannot be applied to several practically important problems. Falin and Templeton's monograph [7] was published in 1997, and summarizes the results of this branch. In their model secondary requests return after some random time, and the case where secondary requests also form a Poisson-process is thoroughly examined.

There are queuing systems, where a request for service can be repeated only after a constant period of time. A typical example can be an automatic redialling device, which again and again attempts to dial the called number after a deterministic time if the line is engaged. This also occurs at airports where planes can start landing upon arrival or have to start a circular manoeuvre, when the runway is used or there are other planes waiting, and they can only put their next request to land after they have completed a full cycle. An application in digital technology is the operation of optical buffers, which will be described later.

Consider an airport where aeroplanes come to land. The airport can serve only one plane, hence if the runway is used or there are other planes waiting to land, the incoming plane has to wait. However, in this system special conditions prevail, which results in great differences from an ordinary queuing system. It is assumed that a plane planning to land approaches the runway in the optimum position, and if it is possible it starts to land on arrival. If the plane is forced to wait then it starts a circular manoeuvre and can put its further requests to land when reaching the original starting point of its trajectory. Completing each full cycle is supposed to take equal time T, thus the possible instants of starting to land can differ from the moment of arrival by integer multiples of T. Because of possible fuel shortage it is natural to use the first-come-first-served (FCFS) discipline.

Although the model stems from a real problem, it certainly is only a simplified version of it, which affects its applicability. For instance, it is assumed without any statistical investigation of real data, that arriving entities form Poisson-processes. Presumably, even if they really do so, the Poisson-processes cannot be homogeneous for almost sure. The FCFS rule is often broken in real life, as well; normally the plane reaching the starting position first is to commence landing. Nevertheless, this simplified model provides exciting tasks to solve, and it can be modified later to fit the real case more precisely.

Originating from this above-mentioned real problem, Lakatos has extensively investigated such queuing systems where the service of a request can be started upon arrival (in case of a free system) or at times differing from it by multiples of the cycle time T (in case of a busy server).

In the basic model [53] entering customers form Poisson-processes of parameter λ, service times are independent, exponentially distributed random variables of parameter μ. Due to the restriction on the start of the services, the functioning of the system is not continuous. There might be idle periods because of the positioning, but the effect of these becomes less as T decreases; in the case $T = 0$ the service becomes continuous, and classic results corresponding to $M/M/1$ systems are received [60]. In [54] service time is uniformly distributed and Pollaczek–Khinchin's formula is derived as a limit case [60].

Kovalenko's followers described the problem in a different way. It was previously proved that waiting times in $GI/G/n$ systems form a Markov-chain [18]. Using this, in [34] the existence of equilibrium is examined from the aspect of the waiting time. While [53] determined the equilibrium distribution, as well as the necessary and sufficient condition of its existence, Koba gives only a sufficient condition, though for a more general case [44]; e. g. in [37, 38] she gives a sufficient condition for the existence of equilibrium of the $M/D/1$ system. Kovalenko generalizes Koba's [34] results, and determines the probability of losing a customer in the $M/G/m$ system when the load is light [49].

In classical retrial systems the first customer (which can be either primary or secondary) is accepted for service by the free server, whilst in the systems investigated in the mentioned articles and the dissertation the service discipline is FIFO (first-in-first-out). These two cases are compared in [46, 70] with the help of a 'cost function' and FIFO rule is proved to be more efficient than that applied by classical retrial systems for sufficiently small cycle-times.

It was Kovalenko's suggestion to generalize the problem for two different types of customers. In the system only one customer of first type can be present, it can only be accepted for service in the case of a free system, whereas in all other cases the requests of such customers are turned down. There is no such restriction on customers of second type; they are serviced immediately or join a queue in case of a busy server. This model can be applied for systems accepting impatient customers who need urgent service, as it is mentioned in [65]. If they cannot get serviced, they leave the system and find another server which is free. I examined this type of system with different continuous distributions in [21, 22, 23, 30], with discrete ones in [26, 27, 29]; simulation results were also included in [30, 11].

Chapters 2 and 3 summerize my results on Lakatos-type queuing systems with two different priority customers and various continuous and discrete service time distributions, respectively.

In Chapter 2 an embedded Markov-chain is defined, whose states are identified with the number of customers in the system at moments just before the service of a new customer begins. For this chain the following probabilities are introduced:

a_{ji}: the probability of appearance of i customers of second type at the service of a j^{th}-type customer ($j = 1, 2$) if at the beginning there is only one customer in the system;

b_i: the probability of appearance of i customers of second type at the service of a second-type customer, if at the beginning of service there are at least two customers in the system;

c_i: the probability of appearance of i customers of second type after free state.

These transition probabilities, as well as the equilibrium probabilities are given through their generating functions.

Theorem 2.1.1 *Consider a queuing system with two types of customers forming Poisson-processes with parameters λ_1 and λ_2. There is no restriction on customers of second type; however, customers of first type may only join the system when it is free (and only one of them can be present at every instant), all other requests of this type are refused. The service of a customer may start at the moment of its arrival (in case of a free system) or at moments differing from it by multiples of cycle time T; and the FIFO (FCFS) rule is obeyed. An embedded Markov-chain is defined, whose states correspond to the number of customers in the system at moments just before starting a service. Service time distributions of customers of either type may be exponential with parameters μ_i or uniform in the intervals $[\alpha_j, \beta_j]$; thus four different cases are to be considered $(i, j = 1, 2)$.*

The matrix of transition probabilities of this chain has the form:

$$\begin{pmatrix} c_0 & c_1 & c_2 & c_3 & \dots \\ a_{20} & a_{21} & a_{22} & a_{23} & \dots \\ 0 & b_0 & b_1 & b_2 & \dots \\ 0 & 0 & b_0 & b_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The elements of the matrix are determined by their generating functions below. The type of service time distribution is indicated in the upper index: $A_j^{\{\exp,uni\}}(z)$ indicates the type of service time distribution of j^{th}-type customers, and $B^{\{\exp,uni\}}(z)$ indicates the type of service time distribution of second-type customers.

$$A_j^{\exp}(z) = \sum_{i=0}^{\infty} a_{1i} z^i = \frac{\mu_j}{\lambda_2 + \mu_j} + \frac{\lambda_2 z}{\lambda_2 + \mu_j} \frac{e^{\lambda_2 (z-1) T}\left(1 - e^{-\mu_j T}\right)}{1 - e^{(\lambda_2 (z-1) - \mu_j) T}},$$

$$A_j^{\mathrm{uni}}(z) = \sum_{i=0}^{\infty} a_{ji} z^i = \frac{e^{-\lambda_2 \alpha_j} - e^{-\lambda_2 \beta_j}}{\lambda_2 \left(\beta_j - \alpha_j\right)} \left(1 - z \frac{1 - e^{-\lambda_2 T}}{1 - e^{-\lambda_2 T z}}\right) + \\ + z \frac{e^{\lambda_2 (z-1) \alpha_j} - e^{\lambda_2 (z-1) \beta_j}}{\lambda_2 \left(\beta_j - \alpha_j\right)} \left(\frac{1 - e^{-\lambda_2 T}}{1 - e^{-\lambda_2 T z}} - \frac{\lambda_2 T}{1 - e^{-\lambda_2 (z-1) T}}\right),$$

$$B^{\exp}(z) = \sum_{i=0}^{\infty} b_i z^i = \frac{1 - e^{\lambda_2 (z-2) T}}{\left(1 - e^{-\lambda_2 T}\right)(2 - z)} - \\ - \frac{\lambda_2}{\lambda_2 (2 - z) + \mu_2} \cdot \frac{1 - e^{\lambda_2 (z-1) T}}{1 - e^{-\lambda_2 T}} \cdot \frac{1 - e^{(\lambda_2 (z-2) - \mu_2) T}}{1 - e^{(\lambda_2 (z-1) - \mu_2) T}},$$

$$B^{\mathrm{uni}}(z) = \sum_{i=0}^{\infty} b_i z^i = \frac{e^{\lambda_2 (z-1) \alpha_2} - e^{\lambda_2 (z-1) \beta_2}}{\left(1 - e^{\lambda_2 (z-1) T}\right)\left(\beta_2 - \alpha_2\right)\left(1 - e^{-\lambda_2 T}\right)} \times \\ \times \left(\frac{1 - e^{\lambda_2 (z-2) T} - e^{\lambda_2 (z-1) T} + e^{\lambda_2 (2z-3) T}}{\lambda_2 (z-2)^2} + T \frac{e^{\lambda_2 (z-2) T} - e^{\lambda_2 (z-1) T}}{z - 2}\right),$$

$$C(z) = \sum_{i=0}^{\infty} c_i z^i = \frac{\lambda_1}{\lambda_1 + \lambda_2} A_1(z) + \frac{\lambda_2}{\lambda_1 + \lambda_2} A_2(z)\,.$$

Theorem 2.1.2 *The generating function of the equilibrium distribution of this chain is:*

$$P(z) = \sum_{i=0}^{\infty} p_i z^i = \frac{p_0\big(zC(z) - B(z)\big) + p_1 z\big(A_2(z) - B(z)\big)}{z - B(z)},$$

where p_0 and p_1 are the first two probabilities of the equilibrium distribution. They are connected with the relation $p_1 = \frac{1-c_0}{a_{20}} p_0$, and

$$p_0 = \frac{1 - B'(1)}{1 - B'(1) + C'(1) + \frac{1-c_0}{a_{20}}\big(A_2'(1) - B'(1)\big)},$$

where

$$
\begin{aligned}
A_j'^{\,\exp}(1) &= \frac{\lambda_2}{\lambda_2+\mu_j}\left(1+\frac{\lambda_2 T}{1-e^{-\mu_j T}}\right),\\
A_j'^{\,\mathrm{uni}}(1) &= \frac{\left(e^{-\lambda_2\beta_j}-e^{-\lambda_2\alpha_j}\right)\left(1+\lambda_2 T-e^{\lambda_2 T}\right)}{\lambda_2\left(\beta_j-\alpha_j\right)\left(1-e^{\lambda_2 T}\right)}+\frac{\lambda_2\left(\alpha_j+\beta_j+T\right)}{2},\\
B'^{\,\exp}(1) &= 1-\frac{\lambda_2 T}{1-e^{-\lambda_2 T}}\left(e^{-\lambda_2 T}-\frac{\lambda_2}{\lambda_2+\mu_2}\cdot\frac{1-e^{-(\lambda_2+\mu_2)T}}{1-e^{-\mu_2 T}}\right),\\
B'^{\,\mathrm{uni}}(1) &= \frac{\lambda_2\left(\alpha_2+\beta_2+T\right)}{2},\\
C'(1) &= \frac{\lambda_1}{\lambda_1+\lambda_2}A_1'(1)+\frac{\lambda_2}{\lambda_1+\lambda_2}A_2'(1)\,.
\end{aligned}
$$

It is also shown that the ergodicity of the process does not depend on customers of first type; its necessary and sufficient condition is stated in the next theorem.

Theorem 2.1.4 *The condition of existence of the ergodic distribution is the fulfilment of one of the following inequalities.*

If the service time of second-type customers is exponentially distributed (regardless of the distribution of service time of first-type customers):

$$
\frac{\lambda_2}{\mu_2} < e^{-\lambda_2 T}\frac{1-e^{-\mu_2 T}}{1-e^{-\lambda_2 T}}. \tag{1.1}
$$

If the service time of second-type customers is uniformly distributed (regardless of the distribution of service time of first-type customers):

$$
\frac{\lambda_2\left(\alpha_2+\beta_2+T\right)}{2} < 1. \tag{1.2}
$$

The dependence on input parameters of condition (1.2) is simple, but it is more complex in (1.1). For the latter case it is more convenient if parameters that guarantee ergodicity are represented in a graph. The shaded areas (the areas under the curves, excluding the curves themselves) in Figure 2.3 show pairs (μ_2, λ_2) suitable for ergodicity at some arbitrary values of the cycle-time (at $T = 0.2$, 0.4, 0.6, 0.8, and 1.0). The graph also displays the limit case $T \to 0$.

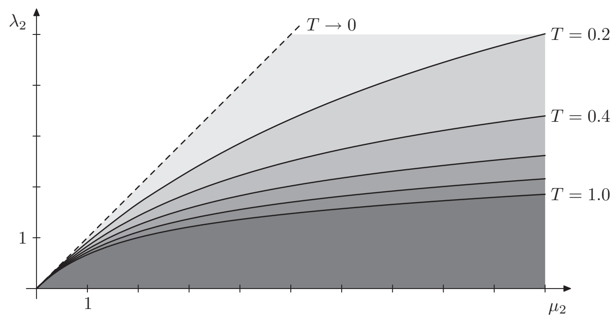


Figure 2.3: *Pairs* (μ_2, λ_2) *suitable for ergodicity*

It can be clearly seen that at fixed capacity of the server (at constant μ_2) the longer the cycle-time is, the lower input intensity the system can tolerate. Additionally, at fixed intensity of the input flow, with longer cycle-time the server has to be of higher performance.

The influence of idle periods — while the system is waiting for the next entity to reach its starting position to be able to start its service — becomes less and less while $T \to 0$. Within this transition conditions of ergodicity (1.1) and (1.2) turn into classical conditions $\frac{1}{\mu_2} < \frac{1}{\lambda_2}$ (see Figure 2.3) and $\frac{\alpha_2+\beta_2}{2} < \frac{1}{\lambda_2}$, i. e. the expectation of the service time must be less than the expectation of the inter-arrival time. In the light of this, (1.2) is even more straightforward: the mean value of the service time increased by the average idle time (on average $\frac{T}{2}$ time is needed for the next customer in the queue to reach the starting position) has to be less than the average inter-arrival time. Unfortunately, (1.1) cannot be interpreted in such a clear probabilistic way.

As a corollary to the theorems, the limit distributions while $T \to 0$ are also given for all combinations of the service time distributions. In a separate section some generalizations are carried out, with no restrictions on the boundaries of the uniform distributions. To validate results, equilibrium probabilities of states 0 and 1 are taken a closer look, examining their dependence on the input parameters.

Unfortunately, most of the models are so complicated that they cannot be treated and solved analytically. In these cases simulation and numerical investigation are useful tools to examine the behaviour of the systems. For instance, the MOSEL tool was developed by scientists of the University of Debrecen (J. Sztrik et al.) and the University of Erlangen (G. Bolch et al.) for this purpose [1, 2, 3, 73, 74, 75]. To underpin my theory, numerical computer experiments are carried out, which complete the investigation of the continuous case.

In the light of technical applications it is important to consider discrete models. In this case cycle-time is divided into n equal time-slices, which form the basis of the discrete distributions. A typical application in digital technology is the use of an optical buffer, which is a device that is capable of temporarily storing light (or rather, data in the form of light). As light cannot be frozen, a typical optical buffer is realized by a single loop, in which data circulate a variable number of times, and thus n can measure the cycle-time in clock cycles. The model was investigated by Rogiest, Laevens et al. in [51, 52, 72]. The involved optical buffers are implemented as a set of $N+1$ Fiber Delay Lines (FDLs), with lengths that are typically a multiple of a basic value D called the granularity. This results in a degenerate waiting room with waiting times 0, D, $2D$, …, ND. The problem is investigated from the point of view of the customers, i. e. their waiting time. Lakatos and myself chose a different approach, the problem is described from the aspect of the server, i. e. the number of waiting customers, which is more significant for determining the number of necessary FDLs. The time elapsed between two arrivals was geometrically distributed and service times of customers were geometric and uniform in [56] and [67], respectively. The two different approaches to describe these systems coincide in the condition of ergodicity and the probability of free state; this was shown in [69]. A numerical investigation was carried out in [11]. I investigated a system with the above-mentioned relative priority property with geometric inter-arrival and service time distributions in [29], and discrete uniform service time distributions in [27].

Chapter 3 is devoted to the above-mentioned discrete, relative-priority Lakatos-type system. The method in this case follows a similar routine as in Chapter 2. Cycle-time T is divided into n equal time-slices. The probability of appearance of a j^{th}-type customer during a certain time-slice is r_j, i. e. inter-arrival times are geometrically distributed with parameters r_j $(j = 1, 2)$. However, unlike with continuously distributed service times, in discrete systems different types of customers do appear during the same time-slice with non-zero probability. There are several ways to deal with this phenomenon, called collision; we opt for three methods to treat it: both of them are refused; first-type customers are accepted, but second-type ones are refused; both of them are accepted, but first-type ones are served first.

The same embedded Markov-chain is defined as in the continuous case, but when applying the third collision treatment method, in addition to previously defined transition probabilities new ones have to be introduced. Let a_{12i} denote the probability of appearance of i customers of second-type at the service of a first-type customer, if the service process started with the simultaneous appearance of customers of both types. Two service time distributions are considered; for one possibility service times are geometrically distributed with parameters q_j, i. e. the service of a j^{th}-type customer continues during a time unit with probability q_j, and terminates with probability $1 - q_j$. The other examined alternative is when service time distributions are uniform in the intervals $[\gamma_j, \delta_j]$, where γ_j and δ_j are multiples of T, i. e. the probability that the service of a j^{th}-type customer is any time units in this interval is $q_j = \frac{T}{n(\delta_j - \gamma_j)}$. All three collision disciplines are investigated with each system. The matrix of transition probabilities is the same as in the continuous case; their generating functions, as well as that of the equilibrium probabilities are determined.

Theorem 3.1.1 *Consider a discrete cyclic-waiting system serving two types of customers in which both inter-arrival time distributions are geometric with parameters r_j $(j = 1, 2)$. Service time distributions of customers of either type may be geometric with parameters q_i or uniform in the intervals $[\gamma_j, \delta_j]$; thus (considering the three collision disciplines) 12 different cases are to be examined $(i, j = 1, 2)$. The service of an entering customer may start imme-*

diately on arrival if the server is free, but in case of a busy server or waiting customers, first-type customers are refused, and second-type customers join the queue. The service of queued customers may start at times differing from their arrival times by multiples of cycle-time T, which is divided into n equal time slices; these form the units of the probability distributions. The states of the corresponding embedded Markov-chain are identified with the number of customers in the system at moments just before starting the service of a customer. The matrix of transition probabilities is identical with the one in the continuous case; their generating functions are given below. The type of service time distribution is indicated in the upper index: $A_j^{\{\text{geo,uni}\}}(z)$ indicates the type of service time distribution of j^{th}-type customers, $A_{12}^{\{\text{geo,uni}\}}(z)$ indicates the type of service time distribution of first-type customers, and $B^{\{\text{geo,uni}\}}(z)$ indicates the type of service time distribution of second-type customers.

$$
\begin{aligned}
A_j^{\text{geo}}(z) = {} & \frac{(1-r_2)(1-q_j)}{1-q_j(1-r_2)} + z\frac{r_2(1-q_j)}{1-q_j(1-r_2)} + \\
& + z\frac{r_2 q_j\left(1-q_j^n\right)(1-r_2+r_2 z)^n}{(1-q_j(1-r_2))\left(1-q_j^n(1-r_2+r_2 z)^n\right)},
\end{aligned}
$$

$$
\begin{aligned}
A_j^{\text{uni}}(z) = {} & \frac{q_j}{r_2}\left[(1-r_2)^{\frac{\gamma_j}{T}n+1} - (1-r_2)^{\frac{\delta_j}{T}n+1}\right] + \\
& + zq_j\left[(1-r_2)^{\frac{\gamma_j}{T}n} - (1-r_2)^{\frac{\delta_j}{T}n}\right] + zq_j\left[(1-r_2)^{\frac{\gamma_j}{T}n} - (1-r_2)^{\frac{\delta_j}{T}n}\right] \times \\
& \times \frac{1-(1-r_2)^n}{r_2(1-r_2)^n}(1-r_2+r_2 z)^n \frac{1-\left(\frac{1-r_2+r_2 z}{1-r_2}\right)^{\frac{\gamma_j}{T}n}}{1-\left(\frac{1-r_2+r_2 z}{1-r_2}\right)^n} + \\
& + zq_j(1-r_2+r_2 z)^n\left[n\frac{(1-r_2+r_2 z)^{\frac{\gamma_j}{T}n} - (1-r_2+r_2 z)^{\frac{\delta_j}{T}n}}{1-(1-r_2+r_2 z)^n} - \right. \\
& \left. - (1-r_2)^{\frac{\delta_j}{T}n}\frac{1-(1-r_2)^n}{r_2(1-r_2)^n}\frac{\left(\frac{1-r_2+r_2 z}{1-r_2}\right)^{\frac{\gamma_j}{T}n} - \left(\frac{1-r_2+r_2 z}{1-r_2}\right)^{\frac{\delta_j}{T}n}}{1-\left(\frac{1-r_2+r_2 z}{1-r_2}\right)^n}\right],
\end{aligned}
$$

$$A_{12}^{geo}(z) = z\frac{(1-q_1^n)\,(1-r_2+r_2z)^n}{1-q_1^n\,(1-r_2+r_2z)^n},$$

$$A_{12}^{uni}(z) = znq_1\frac{(1-r_2+r_2z)^{\left(\frac{\gamma_1}{T}+1\right)n}-(1-r_2+r_2z)^{\left(\frac{\delta_1}{T}+1\right)n}}{1-(1-r_2+r_2z)^n},$$

$$\begin{aligned}B^{geo}(z) = {} & \frac{1-(1-r_2)^n\,(1-r_2+r_2z)^n}{1-(1-r_2)\,(1-r_2+r_2z)}\cdot\frac{r_2\,(1-r_2+r_2z)}{1-(1-r_2)^n}+ \\ & +\frac{1-q_2^n\,(1-r_2)^n\,(1-r_2+r_2z)^n}{1-q_2\,(1-r_2)\,(1-r_2+r_2z)}\times \\ & \times\frac{r_2q_2\,(1-r_2+r_2z)\,((1-r_2+r_2z)^n-1)}{(1-(1-r_2)^n)\,(1-(1-r_2)^n\,(1-r_2+r_2z)^n)},\end{aligned}$$

$$\begin{aligned}B^{uni}(z) = {} & \frac{r_2q_2}{1-(1-r_2^n)}\frac{(1-r_2+r_2z)^{\frac{\gamma_2}{T}n}-(1-r_2+r_2z)^{\frac{\delta_2}{T}n}}{1-(1-r_2+r_2z)^n}\times \\ & \times\Bigg[\Big((1-r_2+r_2z)-(1-r_2+r_2z)^{n+1}\Big)\times \\ & \times\left(\frac{1-(1-r_2)^n\,(1-r_2+r_2z)^n}{(1-(1-r_2)\,(1-r_2+r_2z))^2}--\frac{n\,(1-r_2)^n\,(1-r_2+r_2z)^n}{1-(1-r_2)\,(1-r_2+r_2z)}\right)+ \\ & +n\,(1-r_2+r_2z)^{n+1}\frac{1-(1-r_2)^n\,(1-r_2+r_2z)^n}{1-(1-r_2)\,(1-r_2+r_2z)}\Bigg]\end{aligned}$$

and $C(z)$ depends on collision policies:

I. $C(z) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2}A_1(z)+\frac{r_2(1-r_1)}{r_1+r_2-r_1r_2}A_2(z)+\frac{r_1r_2}{r_1+r_2-r_1r_2},$

II. $C(z) = \frac{r_1}{r_1+r_2-r_1r_2}A_1(z)+\frac{r_2(1-r_1)}{r_1+r_2-r_1r_2}A_2(z)\,,$

III. $C(z) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2}A_1(z)+\frac{r_2(1-r_1)}{r_1+r_2-r_1r_2}A_2(z)+\frac{r_1r_2}{r_1+r_2-r_1r_2}A_{12}(z)\,.$

Theorem 3.1.2 *The generating function of the equilibrium distribution of the defined chain is:*

$$P(z) = \sum_{i=0}^{\infty}p_iz^i = \frac{p_0\big(zC(z)-B(z)\big)+p_1z\big(A_2(z)-B(z)\big)}{z-B(z)},$$

where p_0 and p_1 are the first two probabilities of the equilibrium distribution. They are connected with the relation $p_1 = \frac{1-c_0}{a_{20}} p_0$, and

$$p_0 = \frac{1 - B'(1)}{1 - B'(1) + C'(1) + \frac{1-c_0}{a_{20}} \big(A'_2(1) - B'(1) \big)},$$

where

$$A_j'^{\,geo}(1) = \frac{r_2}{1 - q_j\,(1 - r_2)} \left(1 + \frac{n r_2 q_j}{1 - q_j^n} \right),$$

$$A_j'^{\,uni}(1) = -a_{j0} + \frac{T}{\delta_j - \gamma_j} \left[(1 - r_2)^{\frac{\gamma_j}{T} n} - (1 - r_2)^{\frac{\delta_j}{T} n} \right] \frac{(1 - r_2)^n}{1 - (1 - r_2)^n} + \\ + \frac{n r_2}{T} \frac{\gamma_j + \delta_j + T}{2},$$

$$A_{12}'^{\,geo}(1) = 1 + \frac{n r_2}{1 - q_1^n},$$

$$A_{12}'^{\,uni}(1) = 1 + \frac{n r_2}{T} \frac{\gamma_1 + \delta_1 + T}{2},$$

$$B'^{\,geo}(1) = 1 - \frac{n r_2}{1 - (1 - r_2)^n} \left[(1 - r_2)^n - \frac{r_2 q_2}{1 - q_2^n} \frac{1 - q_2^n\,(1 - r_2)^n}{1 - q_2\,(1 - r_2)} \right],$$

$$B'^{\,uni}(1) = \frac{n r_2}{T} \frac{\gamma_2 + \delta_2 + T}{2},$$

and $C'(1)$ depends on collision policies:

I. $C'(1) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2} A'_1(1) + \frac{r_2(1-r_1)}{r_1+r_2-r_1r_2} A'_2(1)\,,$

II. $C'(1) = \frac{r_1}{r_1+r_2-r_1r_2} A'_1(1) + \frac{r_2(1-r_1)}{r_1+r_2-r_1r_2} A'_2(1)\,,$

III. $C'(1) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2} A'_1(1) + \frac{r_2(1-r_1)}{r_1+r_2-r_1r_2} A'_2(1) + \frac{r_1 r_2}{r_1+r_2-r_1r_2} A'_{12}(1)\,.$

It is also shown that ergodicity of the process depends neither on first-type customers nor on collision policies, and its necessary and sufficient condition is stated in the next theorem.

Theorem 3.1.4 *The condition of existence of the equilibrium distribution is the fulfilment of one of the following inequalities.*

If the service time of second-type customers is geometrically distributed (regardless of the distribution of service time of first-type customers and the applied collision discipline):

$$\frac{r_2 q_2}{1 - q_2^n} \frac{1 - q_2^n\,(1 - r_2)^n}{1 - q_2\,(1 - r_2)} < (1 - r_2)^n\,. \tag{1.3}$$

If the service time of second-type customers is uniformly distributed (regardless of the distribution of service time of first-type customers and the applied collision discipline):

$$\frac{nr_2}{T}\frac{\gamma_2+\delta_2+T}{2}<1. \tag{1.4}$$

The dependence on input parameters of condition (1.4) is simple and can be interpreted easily: considering that $\frac{T}{n}\frac{1}{r_2}$ is the average inter-arrival time, the condition rewritten in the form

$$\frac{\gamma_2+\delta_2}{2}+\frac{T}{2}<\frac{T}{nr_2}$$

necessitates that the average service time increased by the average idle time (on average $\frac{T}{2}$ time is needed for the next customer in the queue to reach the starting position) be less than the average inter-arrival time.

For the more complex geometric relation (1.3) it is more convenient if parameters that guarantee ergodicity are represented in a graph. Numerical examination of inequality (1.3) reveals that for fixed values of n and q_2, the process is ergodic if $r_2 \in (0, r_2^{\max})$. The shaded areas (the areas under the curves, excluding the curves themselves) in Figure 3.1 show pairs (q_2, r_2) suitable for ergodicity at some arbitrary values of n (at $n = 1$, 2, 4, 6, 8, and 10).

It can be seen that in the (somewhat degenerate) case $n = 1$, the condition of ergodicity is $r_2 < 1 - q_2$, which has a clear probabilistic interpretation; namely, that the probability of appearance of a new customer of second-type during a time slice (which is T in this case) has to be less than the probability of completing a service during the same amount of time.

Like in the continuous case, equilibrium probabilities of states 0 and 1 are also investigated. Additionally, the expected value of the queue-length in the examined type of queuing systems is also given.

Theorem 3.2.1 *The expected value of the length of the queues in cyclic-waiting systems serving two types of customers is given by*

$$P'(1)=1-p_0+\frac{(1-p_0)\,B''(1)+p_0\left(C''(1)+\frac{1-c_0}{a_{20}}\left(A_2''(1)-B''(1)\right)\right)}{2\left(1-B'(1)\right)}.$$

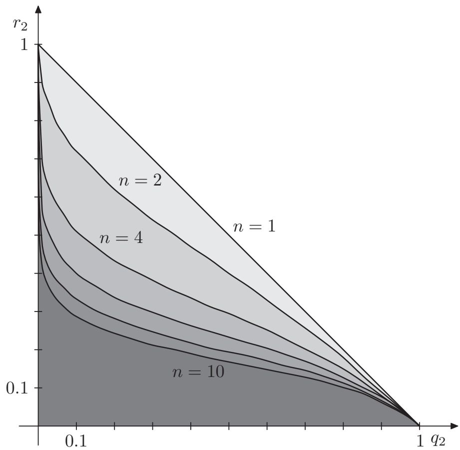


Figure 3.1: *Pairs* (q_2, r_2) *suitable for ergodicity*

Queue-length is explicitly determined in the case of geometric service time distributions, and its dependence on input parameters is analyzed graphically.

Both in the continuous and the discrete case all four possibilities (considering the service time distributions of the two types of customers) have been investigated separately, and presented at different forums [61, 11, 19, 20, 21, 22, 23, 26, 27, 29, 30, 31]. In both chapters the unified theory is given, i. e. generating functions can be combined together as necessary, in accordance with the type of service time distributions of first- and second-type customers. Similarly, the generating function of the equilibrium distribution, as well as the conditions of ergodicity are valid for all cases.

The last chapter of the dissertation is devoted to the investigation of queuing systems with vacation which are closely related to single-server multi-queue polling models. The characteristic feature of polling models is that the server is moving between queues (which possibly requires switchover times), implying that the priority of the queues is dynamically (e.g. cyclically) changing. Some examples are token-passing schemes in local area networks with distributed channel access control, and resource arbitration and load sharing in multiprocessor computers. A vacational system is a queuing system in which the server intermittently spends time away from the queue, perhaps because of a breakdown and repair or because it is serving other jobs. Research on vacation systems has paralleled and frequently overlapped research on polling systems as a system seen from the vantage-point of a particular job class, i. e. it is an example of a vacation system, in which the vacation corresponds to the time spent serving other job classes (plus switchover times). The methods of analysis are similar to those used for polling systems, but because of the focus being on one class, more general results have been obtained. One of the most significant results of the research on vacation systems has been the discovery that the waiting time in the queue in an $M/GI/1$ queue with vacations is distributed as the sum of two independent components, one distributed as the waiting time in the queue in the corresponding $M/GI/1$ queue without vacations and the other as the equilibrium residual time in a vacation. Fuhrmann was apparently the first to identify and prove this decomposition property, which holds under very general conditions. (For example, the length of the vacation can depend on the previous evolution of the queue and the amount of work to be done when the server returns from vacation can depend on the length of the vacation.) Subsequent generalizations were by Doshi, who gave a very general version of the decomposition property.

Chapter 4 deals with the equilibrium distribution of a certain type of $M/G/1$ vacational system, which accepts bulk arrivals. $M/G/1$-type systems are usually analyzed in two different ways. The full description is given by Takács's integro-differential equation (see e. g. [14]), the other applied technique is the Markov-chain method. In the latter case the states of the system are considered at the instants of completing the service of a customer, and the generating function of the equilibrium distribution is determined. The probabilities of the different states are gained through the usual differentiation, although this results in very complicated expressions, especially at higher states, which makes the applicability of the whole method rather difficult.

The mentioned difficulties induced the search for other methods. In [4] and [5] Brière and Chaudhry investigated the $M/G/1$ system with bulk-arrivals and bulk-service, the invertation of the generating function was carried out by comparing the coefficients of the different powers of variable z, and recursive algorithms were given for several particular service time distributions. In [6] a unified approach to the numerical investigation of $M/G/1$ systems is elaborated. A solution is given to the case when the Laplace–Stieltjes-transform of the service time distribution is a rational function; an explicit expression is given with the help of the roots of the characteristic function.

Lakatos applied a different method in [57, 64]. The functioning of the $M/G/1$ system can be described as a regenerative process, in which the end-points of the busy periods are the regeneration points. On the basis of [77], this way equilibrium probabilities can be determined with the help of the expectation of times spent at certain levels and the average length of the busy period. The advantage of this method is that it gives results based only on data received by observing the system directly (the intensity of the input flow, the expectation of the service of a customer, and the probability of appearance of a certain number of customers during the service of another one), it is not necessary to know the distribution function of the service time or its Laplace–Stieltjes-transform.

The novelty is to calculate times spent above level $i-1$ and above level i, based on the idea that they have the same structure; and thus receive a recurrence relation on the expected value of time spent on level i as a difference of the previous ones. [57] describes the ordinary $M/G/1$ system, and the one with vacation at the end of the busy period, whereas [64] deals with vacation at the beginning of the busy period. Bulk-arrivals were considered in [58] and [59]. In [66] Pollaczek–Khinchin's formula is derived from regenerative processes.

In Chapter 4 the results of the previously mentioned systems are generalized, and recursive formulae are given on the equilibrium probabilities of the $M/G/1$ system with vacation at the beginning of the busy period, and which accepts bulk arrivals [24, 25, 28].

Theorem 4.2.1 *In the bulk-arrival $M/G/1$ system with vacation at the beginning of the busy period the ergodic distribution exists if $\varrho < 1$; and equilibrium probabilities are determined by*

$$p_i = \frac{\xi_i'}{\zeta'}, \qquad i = 0, 1, 2, \ldots,$$

where ξ_i' and ζ' are the mean values of time spent on the i^{th} level for a busy period, and the duration of the busy period in the system with vacation, respectively.

The average length of the busy period proves to be $\frac{\eta+\alpha\tau}{1-\varrho}$, where η is the mean length of the vacation, α is the average number of customers arriving in a group, τ is the mean length of a service of a customer, and ϱ is the utilization factor of the system. The mean times spent on certain levels are determined by applying the above method, and are given recursively.

Theorem 4.2.4 *The mean value of time spent on the i^{th} level for a busy period satisfies the recurrence relations*

$$\begin{aligned}
\xi_0' &= \tau, \\
\xi_1' &= \frac{\tau}{c_0} + f_1 (\eta - \tau), \\
\xi_2' &= \xi_2 + (2 - f_1) \left(\frac{\tau}{1-\varrho} - \zeta_1 \right) - \frac{\tau}{c_0} + f_2 (\eta - \tau), \\
&\vdots \\
\xi_i' &= \xi_i + \sum_{k=2}^{i-1} (1 - f_1 - \ldots - f_{i-k}) \, \xi_k + \\
&\quad + (1 - f_1 - \ldots - f_{i-1}) \left(\frac{\tau}{1-\varrho} - \zeta_1 \right) + f_i (\eta - \tau), \qquad (i \geq 3),
\end{aligned}$$

where ξ_i and ζ_i are the mean values of time spent on the i^{th} level, and above it for a busy period in the ordinary (non-vacational) system, respectively; c_k is the probability of appearance of k new groups of customers during the service of another customer; f_k is the probability of k requests being present at the end of vacation.

Chapter 2

Continuous cyclic-waiting retrial systems

The system investigated in this chapter was originally based on a real problem connected with the landing of aeroplanes, but later many other applications emerged which are strongly related to information technology. The situation is going to be described within the initial framing.

Consider an airport where aeroplanes come to land. The airport can serve only one plane, hence if the runway is used or there are other planes waiting to land, the incoming plane has to wait. However, in this system special conditions prevail, which results in great differences from an ordinary queuing system. We assume that a plane planning to land approaches the runway in the optimum position, and if it is possible it starts to land on arrival. If the plane is forced to wait then it starts a circular manoeuvre and can put its further requests to land when reaching the original starting point of its trajectory. We assume that completing each full cycle takes equal time T, thus the possible instants of starting to land can differ from the moment of arrival by integer multiples of T. Because of possible fuel shortage it is natural to use the first-come-first-served (FCFS) discipline.

To put it scientifically, we are investigating a retrial system, where the service of an incoming customer can be started upon arrival if the system is in free state, or if the server is busy or there are other entities waiting, then they join a queue and their service is started at the nearest possible instant differing from the arrival by multiples of cycle-time T. Incoming customers form Poisson-processes, and the FIFO rule is obeyed. Different service time distributions lead to various problems, these were broadly investigated by Lakatos. In [53] service time distribution is exponential, whereas in [54] it is uniform.

It was Kovalenko's suggestion to generalize the problem for two different types of customers. In the system only one customer of first type can be present, it can only be accepted for service in the case of a free system, whereas in all other cases the requests of such customers are turned down. There is no such restriction on customers of second type; they are serviced immediately or join a queue in case of a busy server. In the following sections this system is investigated with continuous distributions, whereas the next chapter is devoted to discrete distributions.

2.1 Theoretical results

Let us consider a queuing system serving two types of customers with their properties described above. Both types of customers form Poisson processes, i. e. interarrival times are independent, identically (exponentially) distributed. Their service time distributions are either exponential or uniform, and not necessarily of the same type; therefore, four different cases are considered.

In conventional queuing systems the service process runs continuously; after having completed the service of a customer, we immediately take the next one. In our system described above, this is not so. Therefore, to elaborate the mathematical description of the system we make the following assumptions. In the system there might be idle periods, when the service of a request is completed, but the next one has not reached its starting position. We consider these periods part of the service time, making the service process continuous in such way. We also make a restriction on the boundaries of the intervals of the uniform distribution: they are multiples of the cycle-time. This assumption does not violate the generality of the theory, but without it formulae are much more complicated. The general case is considered later in Section 2.2.

For the description of the system we are going to use the *embedded Markov-chain* technique. Let us consider the number of customers in the system at moments just before the service of a new customer begins. In other words, if t_k denotes the moment when the service of the k^{th} entity starts, we consider the sequence, whose states correspond to the number of customers at $t_k - 0$, which is denoted by ξ_k. For the sake of definiteness, at $t = 0$ let the system be free. We are going to show that the system $\{\xi_k, k \geq 1\}$ defined this way forms a Markov-chain.

Transitions between states can only occur at the embedded points. Similarly to [33], if ζ_k denotes the number of appearing customers in $[t_{k-1}, t_k]$, then

$$\xi_{k+1} = \begin{cases} \xi_k - 1 + \zeta_{k+1}, & \text{if } \xi_k > 0; \\ \zeta_{k+1}, & \text{if } \xi_k = 0. \end{cases}$$

Random variables corresponding to both times elapsed between the entering of two consecutive customers and service times are independent; and entering customers form Poisson-processes. On the basis of this ζ_{k+1} can be determined, and consequently

$$P(\xi_{k+1} = j \mid \xi_k = i_k, \xi_{k-1} = i_{k-1}, \ldots, \xi_0 = i_0) = P(\xi_{k+1} = j \mid \xi_k = i_k)\,,$$

which means that the process is Markovian.

For this chain we introduce the following transition probabilities:

a_{ji}: the probability of appearance of i customers of second type at the service of a j^{th}-type customer ($j = 1, 2$) if at the beginning there is only one customer in the system;

b_i: the probability of appearance of i customers of second type at the service of a second-type customer, if at the beginning of service there are at least two customers in the system;

c_i: the probability of appearance of i customers of second type after free state.

As the process runs, the busy period can start with a customer of either type. Because of the definition of the chain, the system is still in state 0 while the first customer (i. e. first one in the busy period: any customer finding an idle server and empty queue) is served, as before the start of its service the system was empty. During the service of this customer only second-type customers are accepted for service, they join the queue, and requests of first-type customers are refused. If there are no other requests present, when the service of the next customer (which is obviously second-type) begins, the system turns into state 1, and probabilities of turning into other states from this one are given by a_{2i}. If k customers arrive during the first service, then the chain transforms into state k, as before the start of service of the second customer, k customers are already in the queue. Transitions from states 0 and 1 can be treated similarly, as the number of appearing customers depends only on the service times of the first or second customer. The only difference is that the first customer can be of either type, whereas the second one can only be second-type, and this feature will be observed through the expression of probabilities c_i. The need for introducing probabilities of all other transitions b_i is justified by the fact that in this case the number of customers appearing during a service depends not only on the service time of the actual customer, but also on the mod$T$ inter-arrival time of the next customer, as it will be seen in the proof.

The generating functions of these transition probabilities are, respectively:

$$A_j(z) = \sum_{i=0}^{\infty} a_{ji} z^i \quad (j = 1, 2)\,, \qquad B(z) = \sum_{i=0}^{\infty} b_i z^i, \qquad C(z) = \sum_{i=0}^{\infty} c_i z^i.$$

Summarizing the properties of the system and introducing notations: we consider a queuing system with two types of customers forming Poisson-processes with parameters λ_1 and λ_2. There is no restriction on customers of second type; however, customers of first type may only join the system when it is free (and only one of them can be present at every instant), all other requests of this type are refused. The service of a customer may start at the moment of its arrival (in case of a free system) or at moments differing from it by multiples of cycle time T; and the FIFO (FCFS) rule is obeyed. An embedded Markov-chain is defined, whose states correspond to the number of customers in the system at moments just before starting a service. Service time distributions of customers of either type may be exponential with parameters μ_i or uniform in the intervals $[\alpha_j, \beta_j]$; thus four different cases are to be considered $(i, j = 1, 2)$.

Theorem 2.1.1 *The matrix of transition probabilities of this chain has the form:*

$$\begin{pmatrix} c_0 & c_1 & c_2 & c_3 & \dots \\ a_{20} & a_{21} & a_{22} & a_{23} & \dots \\ 0 & b_0 & b_1 & b_2 & \dots \\ 0 & 0 & b_0 & b_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.1}$$

The elements of the matrix are determined by their generating functions below. The type of service time distribution is indicated in the upper index: $A_j^{\{\exp,\text{uni}\}}(z)$ indicates the type of service time distribution of j^{th}-type customers, and $B^{\{\exp,\text{uni}\}}(z)$ indicates the type of service time distribution of second-type customers.

$$A_j^{\exp}(z) = \sum_{i=0}^{\infty} a_{1i} z^i = \frac{\mu_j}{\lambda_2 + \mu_j} + \frac{\lambda_2 z}{\lambda_2 + \mu_j} \frac{e^{\lambda_2(z-1)T}\left(1 - e^{-\mu_j T}\right)}{1 - e^{(\lambda_2(z-1)-\mu_j)T}}, \tag{2.2}$$

$$\begin{aligned} A_j^{\text{uni}}(z) = \sum_{i=0}^{\infty} a_{ji} z^i &= \frac{e^{-\lambda_2\alpha_j} - e^{-\lambda_2\beta_j}}{\lambda_2\left(\beta_j - \alpha_j\right)} \left(1 - z\frac{1 - e^{-\lambda_2 T}}{1 - e^{-\lambda_2 T z}}\right) + \\ &+ z\frac{e^{\lambda_2(z-1)\alpha_j} - e^{\lambda_2(z-1)\beta_j}}{\lambda_2\left(\beta_j - \alpha_j\right)} \left(\frac{1 - e^{-\lambda_2 T}}{1 - e^{-\lambda_2 T z}} - \frac{\lambda_2 T}{1 - e^{-\lambda_2(z-1)T}}\right), \end{aligned} \tag{2.3}$$

$$\begin{aligned} B^{\exp}(z) = \sum_{i=0}^{\infty} b_i z^i &= \frac{1 - e^{\lambda_2(z-2)T}}{\left(1 - e^{-\lambda_2 T}\right)(2 - z)} - \\ &- \frac{\lambda_2}{\lambda_2(2 - z) + \mu_2} \cdot \frac{1 - e^{\lambda_2(z-1)T}}{1 - e^{-\lambda_2 T}} \cdot \frac{1 - e^{(\lambda_2(z-2)-\mu_2)T}}{1 - e^{(\lambda_2(z-1)-\mu_2)T}}, \end{aligned} \tag{2.4}$$

$$B^{uni}(z) = \sum_{i=0}^{\infty} b_i z^i = \frac{e^{\lambda_2(z-1)\alpha_2} - e^{\lambda_2(z-1)\beta_2}}{\left(1 - e^{\lambda_2(z-1)T}\right)\left(\beta_2 - \alpha_2\right)\left(1 - e^{-\lambda_2 T}\right)} \times$$
$$\times \left(\frac{1 - e^{\lambda_2(z-2)T} - e^{\lambda_2(z-1)T} + e^{\lambda_2(2z-3)T}}{\lambda_2 \left(z-2\right)^2} + T \frac{e^{\lambda_2(z-2)T} - e^{\lambda_2(z-1)T}}{z-2} \right), \quad (2.5)$$

$$C(z) = \sum_{i=0}^{\infty} c_i z^i = \frac{\lambda_1}{\lambda_1 + \lambda_2} A_1(z) + \frac{\lambda_2}{\lambda_1 + \lambda_2} A_2(z) \,. \quad (2.6)$$

It has to be emphasized that all four cases have been investigated separately, and presented at different forums [61, 19, 20, 21, 22, 23, 30, 31]. In the theorem above the unified theory is given, i. e. generating functions (2.2)–(2.5) can be combined together as necessary, in accordance with the type of distribution of first- and second-type customers.

Proof Because of earlier explanation the matrix of transition probabilities is straightforward. However, we underline that probabilities a_{1i} do not appear in it explicitly, as customers of first type can only be accepted when the system is free. These probabilities are represented through probabilities c_i.

For the description of the system we use the *embedded Markov-chain technique*, i. e. we consider the number of customers in the system at moments just before the service of a customer begins. This actually means the number of customers of second type, as first-type customers are refused when the server is busy. We find the transition probabilities for this chain.

We have to distinguish between two cases, depending on whether there are customers waiting or not.

First we consider the case when only one customer of j^{th}-type is present in the system. Let u denote the service time of this customer and v denote the time elapsed between the beginning of its service and the appearance of a new one. In order to be able to determine how many new requests appear in the meantime, we have to know the distribution of the remaining time. For this we calculate the probability of event $u - v < t$. If $F_j(u)$ denotes the distribution function of the service time of j^{th}-type customers,

$$P(0 < u - v < t) = \int_0^\infty \int_v^{v+t} \lambda_2 e^{-\lambda_2 v} \, \mathrm{d}F_j(u) \, \mathrm{d}v,$$

which seems straightforward if service time distribution is exponential:

$$P^{\exp}(0 < u - v < t) = \int_0^\infty \int_v^{v+t} \lambda_2 e^{-\lambda_2 v} \cdot \mu_j e^{-\mu_j u} \, \mathrm{d}u \, \mathrm{d}v = \frac{\lambda_2}{\lambda_2 + \mu_j} \left(1 - e^{-\mu_j t}\right).$$

If service time is uniformly distributed, then two separate calculations have to be carried out, both are illustrated in Figure 2.1.

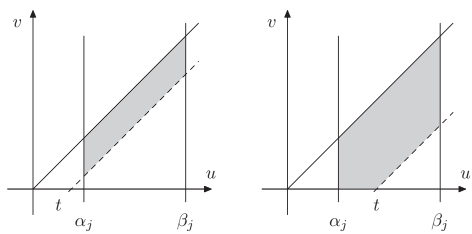


Figure 2.1: *Illustration of integration areas*

If $0 < t \leq \alpha_j$:

$$P^{\text{uni}}(0 < u - v < t) = \int_{\alpha_j}^{\beta_j} \int_{u-t}^{u} \frac{\lambda_2 e^{-\lambda_2 v}}{\beta_j - \alpha_j} \, \mathrm{d}v \, \mathrm{d}u = \frac{1 - e^{\lambda_2 t}}{\lambda_2} \cdot \frac{e^{-\lambda_2 \beta_j} - e^{-\lambda_2 \alpha_j}}{\beta_j - \alpha_j};$$

and if $\alpha_j < t \leq \beta_j$:

$$P^{\text{uni}}(0 < u - v < t) = \int_{\alpha_j}^{\beta_j} \int_0^u - \int_t^{\beta_j} \int_0^{u-t} \frac{\lambda_2 e^{-\lambda_2 v}}{\beta_j - \alpha_j} \, \mathrm{d}v \, \mathrm{d}u =$$
$$= \frac{\lambda_2 (t - \alpha_j) + 1 + e^{-\lambda_2 \beta_j} - e^{-\lambda_2 \alpha_j} - e^{-\lambda_2 (\beta_j - t)}}{\lambda_2 (\beta_j - \alpha_j)}.$$

To get the formula of $A_j(z)$ we have to know how many new requests appear during the service of a customer. We determine the probability of $u-v$ falling between two multiples of T. For customers of exponentially distributed service time:

$$P^{\exp}\big((i-1)\,T < u-v < iT\big) = \frac{\lambda_2}{\lambda_2+\mu_j}\left(e^{-\mu_j(i-1)T} - e^{-\mu_j iT}\right);$$

and for those of uniformly distributed service time,
if $0 < i \le \frac{\alpha_j}{T}$:

$$P^{\text{uni}}\big((i-1)\,T < u-v < iT\big) = \frac{e^{-\lambda_2\alpha_j} - e^{-\lambda_2\beta_j}}{\lambda_2\,(\beta_j-\alpha_j)} e^{\lambda_2 iT}\left(1-e^{-\lambda_2 T}\right);$$

and if $\frac{\alpha_j}{T} < i \le \frac{\beta_j}{T}$:

$$P^{\text{uni}}\big((i-1)\,T < u-v < iT\big) = \frac{\lambda_2 T - e^{\lambda_2(iT-\beta_j)}\left(1-e^{-\lambda_2 T}\right)}{\lambda_2\,(\beta_j-\alpha_j)}.$$

The period between the entry of the second request till the beginning of its service is $\left\lceil\frac{u-v}{T}\right\rceil T$, where $\lceil x\rceil$ denotes the 'upper' integral part of x (the least integer which is not less than x). Considering that $\pi_0 = \int_0^\infty e^{-\lambda_2 x}\,\mathrm{d}F_j(x)$ is the probability that no other requests appear during the service of the present customer, the generating functions of transition probabilities a_{ji} are:

$$A_j(z) = \pi_0 + z\sum_{k=0}^{\infty}\sum_{i=1}^{\infty} P\big((i-1)\,T < u-v < iT\big)\frac{(\lambda_2 iTz)^k}{k!}e^{-\lambda_2 iT}. \tag{2.7}$$

Substituting

$$\pi_0^{\exp} = \int_0^\infty e^{-\lambda_2 x}\cdot\mu_j e^{-\mu_j x}\,\mathrm{d}x = \frac{\mu_j}{\lambda_2+\mu_j},$$

$$\pi_0^{\text{uni}} = \int_{\alpha_j}^{\beta_j} e^{-\lambda_2 x}\cdot\frac{1}{\beta_j-\alpha_j}\,\mathrm{d}x = \frac{e^{-\lambda_2\alpha_j} - e^{-\lambda_2\beta_j}}{\lambda_2\,(\beta_j-\alpha_j)},$$

and the previously calculated probabilities:

$$A_j^{\exp}(z) = \frac{\mu_j}{\lambda_2+\mu_j} + \frac{\lambda_2 z}{\lambda_2+\mu_j}\sum_{k=0}^{\infty}\sum_{i=1}^{\infty}\left(e^{-\mu_j(i-1)T} - e^{-\mu_j iT}\right)\frac{(\lambda_2 iTz)^k}{k!}e^{-\lambda_2 iT},$$

$$
\begin{aligned}
A_j^{\mathrm{uni}}(z) = {} & \frac{e^{-\lambda_2\alpha_j} - e^{-\lambda_2\beta_j}}{\lambda_2\left(\beta_j - \alpha_j\right)} + \frac{z}{\lambda_2\left(\beta_j - \alpha_j\right)} \times \\
& \times \sum_{k=0}^{\infty} \left[\sum_{i=1}^{\frac{\alpha_j}{T}} \left(e^{-\lambda_2\alpha_j} - e^{-\lambda_2\beta_j}\right) e^{\lambda_2 iT} \left(1 - e^{-\lambda_2 T}\right) \frac{\left(\lambda_2 iTz\right)^k}{k!} e^{-\lambda_2 iT} + \right. \\
& \left. + \sum_{i=\frac{\alpha_j}{T}+1}^{\frac{\beta_j}{T}} \left(\lambda_2 T - e^{\lambda_2\left(iT-\beta_j\right)} \left(1 - e^{-\lambda_2 T}\right)\right) \frac{\left(\lambda_2 iTz\right)^k}{k!} e^{-\lambda_2 iT} \right].
\end{aligned}
$$

Expanding sums we get (2.2) and (2.3). The busy period can start with a customer of j^{th} type with probability $\frac{\lambda_j}{\lambda_1+\lambda_2}$, this explains (2.6).

Now we are going to determine the transition probabilities of all other states. In this case, at the instant when the service of the actual request begins, the next one is already present, as well. Let $x = u - \left\lfloor \frac{u}{T} \right\rfloor T$ and y mean the deviation of arrival times mod T. It can easily be seen that y has truncated exponential distribution with distribution function $\frac{1-e^{-\lambda_2 y}}{1-e^{-\lambda_2 T}}$. The period between the starting moments of services of two consecutive requests is

$$
t_0 = \begin{cases} \left\lfloor \frac{u}{T} \right\rfloor T + y, & \text{if } x \le y; \\ \left(\left\lfloor \frac{u}{T} \right\rfloor + 1\right) T + y, & \text{if } x > y. \end{cases}
$$

This is explained in Figure 2.2.

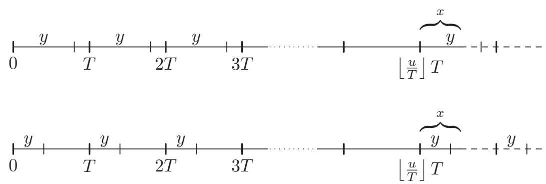


Figure 2.2: *Explanation of t_0*

Let us divide the service time into intervals of length T and fix y. Each such interval is divided into two parts by y (the first part has length y, the second part $T - y$). The probability of appearance of k requests during the investigated period (in the two cases regarding the relation of x and y) is $\frac{(\lambda_2 t_0)^k}{k!} e^{-\lambda_2 t_0}$. Let $\left\lfloor \frac{u}{T} \right\rfloor = i$ and ξ be a random variable denoting the number of requests appearing during the investigated period. The generating function of the number of requests entering the system provided that the mod T interarrival time equals y is:

$$E\left(z^{\xi} \mid y\right) = \sum_{k=0}^{\infty}\sum_{i=0}^{\infty}\left[\int\limits_{iT}^{iT+y}\frac{\left[\lambda_2\left(iT+y\right)z\right]^k}{k!}e^{-\lambda_2(iT+y)}\,\mathrm{d}F_2(u)+\right.$$
$$\left.+\int\limits_{iT+y}^{(i+1)T}\frac{\left[\lambda_2\left(\left(i+1\right)T+y\right)z\right]^k}{k!}e^{-\lambda_2((i+1)T+y)}\,\mathrm{d}F_2(u)\right]. \quad (2.8)$$

If the service time distribution of 2^{nd}-type customers is exponential, then

$$E^{\exp}\left(z^{\xi} \mid y\right) = \sum_{k=0}^{\infty}\sum_{i=0}^{\infty}\left[\int\limits_{iT}^{iT+y}\frac{\left[\lambda_2\left(iT+y\right)z\right]^k}{k!}e^{-\lambda_2(iT+y)}\mu_2 e^{-\mu_2 u}\mathrm{d}u+\right.$$
$$\left.+\int\limits_{iT+y}^{(i+1)T}\frac{\left[\lambda_2\left(\left(i+1\right)T+y\right)z\right]^k}{k!}e^{-\lambda_2((i+1)T+y)}\mu_2 e^{-\mu_2 u}\mathrm{d}u\right]=$$
$$= e^{\lambda_2(z-1)y}\left(1-e^{-\mu_2 y}\frac{1-e^{\lambda_2(z-1)T}}{1-e^{(\lambda_2(z-1)-\mu_2)T}}\right);$$

and if it is uniform, then

$$E^{\mathrm{uni}}\left(z^{\xi} \mid y\right) = \sum_{k=0}^{\infty}\sum_{i=\frac{\alpha_2}{T}}^{\frac{\beta_2}{T}-1}\left[\int\limits_{iT}^{iT+y}\frac{\left[\lambda_2\left(iT+y\right)z\right]^k}{k!}e^{-\lambda_2(iT+y)}\frac{\mathrm{d}u}{\beta_2-\alpha_2}+\right.$$
$$\left.+\int\limits_{iT+y}^{(i+1)T}\frac{\left[\lambda_2\left(\left(i+1\right)T+y\right)z\right]^k}{k!}e^{-\lambda_2((i+1)T+y)}\frac{\mathrm{d}u}{\beta_2-\alpha_2}\right]=$$
$$=\frac{e^{\lambda_2(z-1)\alpha_2}-e^{\lambda_2(z-1)\beta_2}}{1-e^{\lambda_2(z-1)T}}\cdot\frac{e^{\lambda_2(z-1)y}}{\beta_2-\alpha_2}\left(y+\left(T-y\right)e^{\lambda_2(z-1)T}\right).$$

Multiplying these expressions by $\frac{\lambda_2 e^{-\lambda_2 y}}{1-e^{-\lambda_2 T}}$ and integrating with respect to y from 0 to T, we finally obtain the generating functions (2.4) and (2.5) of transition probabilities b_i. ∎

Theorem 2.1.2 *The generating function of the equilibrium distribution of this chain is:*

$$P(z) = \sum_{i=0}^{\infty} p_i z^i = \frac{p_0\big(zC(z) - B(z)\big) + p_1 z\big(A_2(z) - B(z)\big)}{z - B(z)}, \tag{2.9}$$

where p_0 and p_1 are the first two probabilities of the equilibrium distribution. They are connected with the relation $p_1 = \frac{1-c_0}{a_{20}} p_0$, and

$$p_0 = \frac{1 - B'(1)}{1 - B'(1) + C'(1) + \frac{1-c_0}{a_{20}}\big(A_2'(1) - B'(1)\big)}, \tag{2.10}$$

where

$$A_j'^{\,\exp}(1) = \frac{\lambda_2}{\lambda_2 + \mu_j}\left(1 + \frac{\lambda_2 T}{1 - e^{-\mu_j T}}\right), \tag{2.11a}$$

$$A_j'^{\,\mathrm{uni}}(1) = \frac{\left(e^{-\lambda_2\beta_j} - e^{-\lambda_2\alpha_j}\right)\left(1 + \lambda_2 T - e^{\lambda_2 T}\right)}{\lambda_2\left(\beta_j - \alpha_j\right)\left(1 - e^{\lambda_2 T}\right)} + \frac{\lambda_2\left(\alpha_j + \beta_j + T\right)}{2}, \tag{2.11b}$$

$$B'^{\,\exp}(1) = 1 - \frac{\lambda_2 T}{1 - e^{-\lambda_2 T}}\left(e^{-\lambda_2 T} - \frac{\lambda_2}{\lambda_2 + \mu_2} \cdot \frac{1 - e^{-(\lambda_2+\mu_2)T}}{1 - e^{-\mu_2 T}}\right), \tag{2.11c}$$

$$B'^{\,\mathrm{uni}}(1) = \frac{\lambda_2\left(\alpha_2 + \beta_2 + T\right)}{2}, \tag{2.11d}$$

$$C'(1) = \frac{\lambda_1}{\lambda_1 + \lambda_2} A_1'(1) + \frac{\lambda_2}{\lambda_1 + \lambda_2} A_2'(1)\,. \tag{2.11e}$$

Proof The matrix of transition probabilities has the form (2.1). With the help of this we can determine the probabilities of the ergodic distribution denoted by p_l. They satisfy the equations

$$\begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 & \dots \\ a_{20} & a_{21} & a_{22} & a_{23} & \dots \\ 0 & b_0 & b_1 & b_2 & \dots \\ 0 & 0 & b_0 & b_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}^T \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \\ \vdots \end{pmatrix},$$

i. e.

$$p_0 = p_0 c_0 + p_1 a_{20}, \tag{2.12a}$$

$$p_l = \sum_{k=2}^{l+1} p_k b_{l-k+1} + p_0 c_l + p_1 a_{2l} \qquad (l \geq 1)\,, \tag{2.12b}$$

from which we receive the following expression for the generating function:

$$
\begin{aligned}
P(z) = \sum_{l=0}^{\infty} p_l z^l &= p_0 C(z) + p_1 A_2(z) + \sum_{l=1}^{\infty} \sum_{k=2}^{l+1} p_k b_{l-k+1} z^l = \\
&= p_0 C(z) + p_1 A_2(z) + \sum_{k=2}^{\infty} \sum_{l=k-1}^{\infty} p_k b_{l-k+1} z^{l-k+1} z^{k-1} = \\
&= p_0 C(z) + p_1 A_2(z) + B(z) \left(\frac{P(z)}{z} - \frac{p_0}{z} - p_1 \right),
\end{aligned}
$$

which yields (2.9). From (2.12a)

$$
p_1 = \frac{1 - c_0}{a_{20}} p_0. \tag{2.13}
$$

To determine p_0 condition $P(1) = 1$ is used. From this (2.10) is received, where $k = \frac{1-c_0}{a_{20}}$ is a constant factor between the first two probabilities of the ergodic distribution.

To get p_0 as the function of the input parameters, the derivatives of the generating functions have to be determined, and their limits have to be taken as $z \to 1$ applying L'Hôspital's rule. After some tedious calculations (2.11) are received. ■

Lemma 2.1.3 *Expression $C'(1) + k\big(A_2'(1) - B'(1)\big)$ is always positive for any values of the parameters and for any combinations of the service time distributions.*

PROOF As constants $k = \frac{1-c_0}{a_{20}}$ and $C'(1)$ depend on the service time distributions of both type customers, we are going to discuss the four cases separately.

In the following list the explicit values of constants $k^{\mathrm{d}_1,\mathrm{d}_2}$ are given, where d_1 and d_2 in the upper index denote the service time distributions of first- and second-type customers, respectively.

$$
\begin{aligned}
k^{\text{exp,exp}} &= \frac{\lambda_1\lambda_2^2+\lambda_2^3+\lambda_2^2\mu_1+\lambda_1\lambda_2\mu_2}{\mu_2\left(\lambda_1+\lambda_2\right)\left(\lambda_2+\mu_1\right)},\\
k^{\text{exp,uni}} &= \frac{\lambda_2^2\left(\lambda_1+\lambda_2+\mu_1\right)}{\left(\lambda_1+\lambda_2\right)\left(\lambda_2+\mu_1\right)}\cdot\frac{\beta_2-\alpha_2}{e^{-\lambda_2\alpha_2}-e^{-\lambda_2\beta_2}}-\frac{\lambda_2}{\lambda_1+\lambda_2},\\
k^{\text{uni,exp}} &= \frac{\lambda_2\left(\lambda_1\lambda_2+\lambda_2^2+\lambda_1\mu_2\right)}{\left(\lambda_1+\lambda_2\right)\left(\lambda_2+\mu_2\right)}-\frac{\lambda_1}{\lambda_1+\lambda_2}\cdot\frac{e^{-\lambda_2\alpha_1}-e^{-\lambda_2\beta_1}}{\beta_1-\alpha_1},\\
k^{\text{uni,uni}} &= \lambda_2\frac{\beta_2-\alpha_2}{e^{-\lambda_2\alpha_2}-e^{-\lambda_2\beta_2}}-\frac{\lambda_2}{\lambda_1+\lambda_2}-\\
&\qquad -\frac{\lambda_1}{\lambda_1+\lambda_2}\cdot\frac{e^{-\lambda_2\alpha_1}-e^{-\lambda_2\beta_1}}{\beta_1-\alpha_1}\cdot\frac{\beta_2-\alpha_2}{e^{-\lambda_2\alpha_2}-e^{-\lambda_2\beta_2}}.
\end{aligned}
$$

If the service time distributions of customers of both types are exponential, then

$$
\begin{aligned}
&C'(1)+k\big(A_2'(1)-B'(1)\big)=\\
&\quad=\frac{\lambda_2^2T}{\lambda_1+\lambda_2}\left(\frac{\lambda_1}{\left(\lambda_2+\mu_1\right)\left(1-e^{-\mu_1T}\right)}+\frac{\lambda_2}{\left(\lambda_2+\mu_2\right)\left(1-e^{-\mu_2T}\right)}+\right.\\
&\qquad\left.+\frac{\lambda_1\lambda_2+\lambda_2^2+\lambda_2\mu_1+\lambda_1\mu_2}{\left(\lambda_2+\mu_1\right)\left(\lambda_2+\mu_2\right)}\cdot\frac{e^{-\lambda_2T}}{1-e^{-\lambda_2T}}\right),
\end{aligned}
$$

which is clearly positive.

If the service time distribution of first-type customers is exponential, and that of second-type ones is uniform, then

$$
\begin{aligned}
&C'(1)+k\big(A_2'(1)-B'(1)\big)=\\
&\quad=\frac{\lambda_1\lambda_2^2T}{\left(\lambda_1+\lambda_2\right)\left(\lambda_2+\mu_1\right)}\left(\frac{1}{1-e^{-\mu_1T}}+\frac{1}{e^{\lambda_2T}-1}\right)+\\
&\qquad+\frac{\lambda_2}{\lambda_1+\lambda_2}\left(\frac{\lambda_2\left(\alpha_2+\beta_2\right)}{2}+\frac{1+\lambda_2T-e^{\lambda_2T}}{e^{\lambda_2T}-1}+\frac{\lambda_2T}{2}\right).
\end{aligned}
$$

Substituting e^{λ_2T} in the numerator by its McLaurin-series and combining the last two terms one can see that all terms are positive:

$$
\begin{aligned}
&C'(1) + k\big(A_2'(1) - B'(1)\big) = \\
&\qquad = \frac{\lambda_1\lambda_2^2 T}{(\lambda_1+\lambda_2)\,(\lambda_2+\mu_1)}\left(\frac{1}{1-e^{-\mu_1 T}} + \frac{1}{e^{\lambda_2 T}-1}\right) + \\
&\qquad + \frac{\lambda_2}{\lambda_1+\lambda_2}\left(\frac{\lambda_2\,(\alpha_2+\beta_2)}{2} + \frac{1}{e^{\lambda_2 T}-1}\sum_{k=2}^{\infty}\frac{(\lambda_2 T)^k (k-2)}{2k!}\right) > 0.
\end{aligned}
$$

If the service time distribution of first-type customers is uniform, and that of second-type ones is exponential, then

$$
\begin{aligned}
&C'(1) + k\big(A_2'(1) - B'(1)\big) = \\
&\qquad = \frac{\lambda_2^3 T}{(\lambda_1+\lambda_2)\,(\lambda_2+\mu_2)}\left(\frac{1}{1-e^{-\mu_2 T}} + \frac{1}{e^{\lambda_2 T}-1}\right) + \\
&\qquad + \frac{\lambda_1}{\lambda_1+\lambda_2}\left(\frac{\lambda_2\,(\alpha_1+\beta_1)}{2} + \frac{1+\lambda_2 T - e^{\lambda_2 T}}{e^{\lambda_2 T}-1} + \frac{\lambda_2 T}{2}\right).
\end{aligned}
$$

Substituting $e^{\lambda_2 T}$ in the numerator by its McLaurin-series and combining the last two terms one can see that all terms are positive:

$$
\begin{aligned}
&C'(1) + k\big(A_2'(1) - B'(1)\big) = \\
&\qquad = \frac{\lambda_2^3 T}{(\lambda_1+\lambda_2)\,(\lambda_2+\mu_2)}\left(\frac{1}{1-e^{-\mu_2 T}} + \frac{1}{e^{\lambda_2 T}-1}\right) + \\
&\qquad + \frac{\lambda_1}{\lambda_1+\lambda_2}\left(\frac{\lambda_2\,(\alpha_1+\beta_1)}{2} + \frac{1}{e^{\lambda_2 T}-1}\sum_{k=2}^{\infty}\frac{(\lambda_2 T)^k (k-2)}{2k!}\right) > 0.
\end{aligned}
$$

Finally, if the service time distributions of customers of both types are uniform, then

$$
\begin{aligned}
C'(1) + k\big(A_2'(1) - B'(1)\big) &= \frac{1+\lambda_2 T - e^{\lambda_2 T}}{e^{\lambda_2 T}-1} + \frac{\lambda_2 T}{2} + \\
&\qquad + \frac{\lambda_2}{2}\cdot\frac{\lambda_1(\alpha_1+\beta_1)+\lambda_2(\alpha_2+\beta_2)}{\lambda_1+\lambda_2}.
\end{aligned}
$$

Substituting $e^{\lambda_2 T}$ in the numerator by its McLaurin-series and combining the first two terms one can see that it is positive:

$$C'(1) + k\big(A_2'(1) - B'(1)\big) = \frac{1}{e^{\lambda_2 T} - 1} \sum_{k=2}^{\infty} \frac{(\lambda_2 T)^k (k-2)}{2k!} + \\ + \frac{\lambda_2}{2} \cdot \frac{\lambda_1(\alpha_1 + \beta_1) + \lambda_2(\alpha_2 + \beta_2)}{\lambda_1 + \lambda_2} > 0. \quad \blacksquare$$

Theorem 2.1.4 *The condition of existence of the ergodic distribution is the fulfilment of one of the following inequalities.*

If the service time of second-type customers is exponentially distributed (regardless of the distribution of service time of first-type customers):

$$\frac{\lambda_2}{\mu_2} < e^{-\lambda_2 T} \frac{1 - e^{-\mu_2 T}}{1 - e^{-\lambda_2 T}}. \tag{2.14}$$

If the service time of second-type customers is uniformly distributed (regardless of the distribution of service time of first-type customers):

$$\frac{\lambda_2 \left(\alpha_2 + \beta_2 + T\right)}{2} < 1. \tag{2.15}$$

Proof As the embedded Markov-chain is irreducible and aperiodic, the process is ergodic iff $0 < p_0 < 1$. Applying Theorem 2.1.2 and Lemma 2.1.3, $p_0 = \frac{1-B'(1)}{1-B'(1)+K}$, where K is a positive constant. Thus, the condition simplifies into

$$1 - B'(1) > 0,$$

which – together with (2.11c) and (2.11d) – gives (2.14) and (2.15). ■

One remarkable thing about (2.14) and (2.15) is that they do not depend on first-type customers, i. e. they have no effect on the ergodicity of the process.

The dependence on input parameters of condition (2.15) is simple, but it is more complex in (2.14). For the latter case it is more convenient if parameters that guarantee ergodicity are represented in a graph. The shaded areas (the areas under the curves, excluding the curves themselves) in Figure 2.3 show pairs (μ_2, λ_2) suitable for ergodicity at some arbitrary values of the

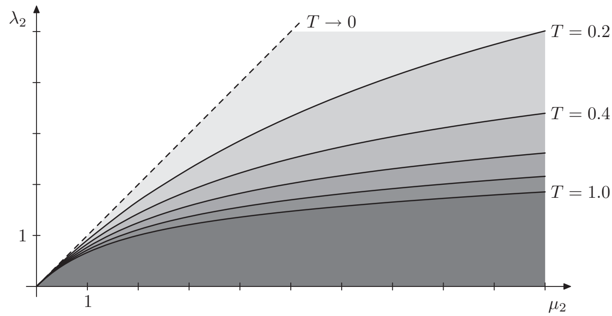


Figure 2.3: *Pairs (μ_2, λ_2) suitable for ergodicity*

cycle-time (at $T = 0.2$, 0.4, 0.6, 0.8, and 1.0). The graph also displays the limit case $T \to 0$.

It can be clearly seen that at fixed capacity of the server (at constant μ_2) the longer the cycle-time is, the lower input intensity the system can tolerate. Additionally, at fixed intensity of the input flow, with longer cycle-time the server has to be of higher performance.

The influence of idle periods — while the system is waiting for the next entity to reach its starting position to be able to start its service — becomes less and less while $T \to 0$. Within this transition conditions of ergodicity (2.14) and (2.15) turn into classical conditions $\frac{1}{\mu_2} < \frac{1}{\lambda_2}$ (see Figure 2.3) and $\frac{\alpha_2+\beta_2}{2} < \frac{1}{\lambda_2}$, i. e. the expectation of the service time must be less than the expectation of the inter-arrival time. In the light of this, (2.15) is even more straightforward: the mean value of the service time increased by the average idle time (on average $\frac{T}{2}$ time is needed for the next customer in the queue to reach the starting position) has to be less than the average inter-arrival time. Unfortunately, (2.14) cannot be interpreted in such a clear probabilistic way.

It is also useful to determine the generating functions of equilibrium distributions during the mentioned $T \to 0$ transition.

Corollary 2.1.5 *The limit distributions $P^*(z)$ as $T \to 0$ are given by the following formulae.*

If service times are exponentially distributed for both types of customers, then

$$P^*(z) = p_0^* \cdot \frac{\lambda_1\mu_1\mu_2 + \lambda_2\mu_1\left(\mu_2 - \lambda_1 z\right) + \lambda_2\mu_2\left(\lambda_1 - \lambda_2\left(z-1\right)\right)}{\left(\lambda_1+\lambda_2\right)\left(\mu_2 - \lambda_2 z\right)\left(\mu_1 - \lambda_2\left(z-1\right)\right)},$$

where

$$p_0^* = \frac{1 - \frac{\lambda_2}{\mu_2}}{1 + \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}\left(\frac{1}{\mu_1} - \frac{1}{\mu_2}\right)}.$$

If service time distribution of first-type customers is exponential and that of second-type customers is uniform, then

$$\begin{aligned}P^*(z) = {} & \frac{p_0^*\left(\beta_2-\alpha_2\right)}{\left(\lambda_1+\lambda_2\right)\left(\lambda_2\left(\beta_2-\alpha_2\right)\left(z-z^2\right) - e^{\lambda_2(z-1)\alpha_2} + e^{\lambda_2(z-1)\beta_2}\right)}\times \\ & \times\left(\frac{\lambda_1\lambda_2\mu_1 z\left(1-z\right)}{\mu_1+\lambda_2\left(1-z\right)} + \frac{\left(\lambda_2\left(z-1\right)-\lambda_1\right)\left(e^{\lambda_2(z-1)\alpha_2} - e^{\lambda_2(z-1)\beta_2}\right)}{\beta_2-\alpha_2}\right),\end{aligned}$$

where

$$p_0^* = \frac{1 - \frac{\lambda_2(\alpha_2+\beta_2)}{2}}{1 + \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}\left(\frac{1}{\mu_1} - \frac{\alpha_2+\beta_2}{2}\right)}.$$

If service time distribution of first-type customers is uniform and that of second-type customers is exponential, then

$$\begin{aligned}P^*(z) = {} & \frac{p_0^*}{\lambda_2\left(\lambda_1+\lambda_2\right)\left(\beta_1-\alpha_1\right)\left(\mu_2-\lambda_2 z\right)}\times \\ & \times\left[\frac{\lambda_1 z\left(\mu_2-\lambda_2\left(z-1\right)\right)\left(e^{\lambda_2(z-1)\beta_1} - e^{\lambda_2(z-1)\alpha_1}\right)}{\left(z-1\right)^2}- \right. \\ & \left. -\frac{\mu_2\lambda_2\left(\beta_1-\alpha_1\right)\left(\lambda_1-\lambda_2\left(z-1\right)\right)}{z-1}\right],\end{aligned}$$

where

$$p_0^* = \frac{1 - \frac{\lambda_2}{\mu_2}}{1 + \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}\left(\frac{\alpha_1+\beta_1}{2} - \frac{1}{\mu_2}\right)}.$$

If service times are uniformly distributed for both types of customers, *then*

$$P^*(z) = \frac{p_0^*\left(\beta_2-\alpha_2\right)}{\left(\lambda_1+\lambda_2\right)\left(\lambda_2\left(\beta_2-\alpha_2\right)\left(z-z^2\right)-e^{\lambda_2(z-1)\alpha_2}+e^{\lambda_2(z-1)\beta_2}\right)}\times$$
$$\times\left(\frac{\lambda_1\left(e^{\lambda_2(z-1)\alpha_1}-e^{\lambda_2(z-1)\beta_1}\right)}{\beta_1-\alpha_1}-\frac{\left(\lambda_2\left(z-1\right)-\lambda_1\right)\left(e^{\lambda_2(z-1)\alpha_2}-e^{\lambda_2(z-1)\beta_2}\right)}{\beta_2-\alpha_2}\right),$$

where

$$p_0^* = \frac{1-\frac{\lambda_2(\alpha_2+\beta_2)}{2}}{1+\frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}\left(\frac{\alpha_1+\beta_1}{2}-\frac{\alpha_2+\beta_2}{2}\right)}.$$

2.2 Amendments

In the previous section some assumptions were taken if service time distributions were uniform, which seemed unnecessary. It was reasonable to suppose that the boundaries of the uniform distributions were multiples of the cycle-time to illustrate our theory easier, to carry out calculations simpler and to take a unified approach with the other distribution considered. However, generating functions in Theorem 2.1.1 are only valid if this assumption is taken. Now we make amendments to generalize formulae.

During the course of determining generating functions we take the same steps as in the proof of Theorem 2.1.1.

In the most general case, when α_j or β_j are not multiples of cycle-time T, several different possibilities have to be considered when calculating the probability of $u-v$ falling between two multiples of T.

If there is at least one integer multiple of T in the interval $(\alpha_j, \beta_j]$ (i. e. $\left\lfloor\frac{\alpha_j}{T}\right\rfloor < \left\lfloor\frac{\beta_j}{T}\right\rfloor$), then
if $0 < i \leq \left\lfloor\frac{\alpha_j}{T}\right\rfloor$:

$$P^{\text{uni}}\Big(\left(i-1\right)T < u-v < iT\Big) = \frac{e^{-\lambda_2\alpha_j}-e^{-\lambda_2\beta_j}}{\lambda_2\left(\beta_j-\alpha_j\right)}e^{\lambda_2 iT}\left(1-e^{-\lambda_2 T}\right);$$

if $i = \left\lfloor\frac{\alpha_j}{T}\right\rfloor+1$:

$$P^{\text{uni}}\Big(\left(i-1\right)T < u-v < iT\Big) =$$
$$= \frac{\lambda_2\left(iT-\alpha_j\right)+1-e^{\lambda_2((i-1)T-\beta_j)}\left(e^{\lambda_2 T}-1\right)-e^{\lambda_2((i-1)T-\alpha_j)}}{\lambda_2\left(\beta_j-\alpha_j\right)};$$

if $\left\lfloor \frac{\alpha_j}{T} \right\rfloor < i \le \left\lfloor \frac{\beta_j}{T} \right\rfloor$:

$$P^{\mathrm{uni}}\big(\,(i-1)\,T < u-v < iT\big) = \frac{\lambda_2 T - e^{\lambda_2 (iT-\beta_j)}\left(1-e^{-\lambda_2 T}\right)}{\lambda_2\left(\beta_j-\alpha_j\right)};$$

and if $i = \left\lfloor \frac{\beta_j}{T} \right\rfloor + 1$:

$$P^{\mathrm{uni}}\big(\,(i-1)\,T < u-v < iT\big) = \frac{\lambda_2\left(\beta_j-(i-1)\,T\right)+1-e^{\lambda_2((i-1)T-\beta_j)}}{\lambda_2\left(\beta_j-\alpha_j\right)}.$$

In the case when there is no integer multiple of T in the interval $(\alpha_j, \beta_j]$ (i. e. $\left\lfloor \frac{\alpha_j}{T} \right\rfloor = \left\lfloor \frac{\beta_j}{T} \right\rfloor$), then
if $0 < i \le \left\lfloor \frac{\alpha_j}{T} \right\rfloor$:

$$P^{\mathrm{uni}}\big(\,(i-1)\,T < u-v < iT\big) = \frac{e^{-\lambda_2\alpha_j}-e^{-\lambda_2\beta_j}}{\lambda_2\left(\beta_j-\alpha_j\right)} e^{\lambda_2 iT}\left(1-e^{-\lambda_2 T}\right);$$

if $i = \left\lfloor \frac{\alpha_j}{T} \right\rfloor + 1$:

$$P^{\mathrm{uni}}\big(\,(i-1)\,T < u-v < iT\big) = 1 - e^{\lambda_2(i-1)T}\cdot\frac{e^{-\lambda_2\alpha_j}-e^{-\lambda_2\beta_j}}{\lambda_2\left(\beta_j-\alpha_j\right)}.$$

We are going to determine $A_j(z)$ according to (2.7), but carry out calculations seperately, regarding the different cases of $P^{\mathrm{uni}}\big(\,(i-1)\,T < u-v < iT\big)$.

If there is at least one integer multiple of T in the interval $(\alpha_j, \beta_j]$ (i. e. $\left\lfloor \frac{\alpha_j}{T} \right\rfloor < \left\lfloor \frac{\beta_j}{T} \right\rfloor$), then

$$\begin{aligned} A^{\mathrm{uni}}_{j,1}(z) &= z\sum_{k=0}^{\infty}\sum_{i=1}^{\left\lfloor \frac{\alpha_j}{T} \right\rfloor}\frac{e^{-\lambda_2\alpha_j}-e^{-\lambda_2\beta_j}}{\lambda_2\left(\beta_j-\alpha_j\right)}e^{\lambda_2 iT}\left(1-e^{-\lambda_2 T}\right)\frac{\left(\lambda_2 iTz\right)^k}{k!}e^{-\lambda_2 iT} = \\ &= z\frac{e^{-\lambda_2\alpha_j}-e^{-\lambda_2\beta_j}}{\lambda_2\left(\beta_j-\alpha_j\right)}\cdot\frac{1-e^{-\lambda_2 T}}{1-e^{-\lambda_2 Tz}}\left(e^{\lambda_2\left\lfloor \frac{\alpha_j}{T} \right\rfloor Tz}-1\right); \end{aligned}$$

$$\begin{aligned} A^{\mathrm{uni}}_{j,2}(z) = \frac{z}{\lambda_2\left(\beta_j-\alpha_j\right)}\Bigg[&\left(\lambda_2\left\lfloor \frac{\alpha_j}{T} \right\rfloor T + \lambda_2\left(T-\alpha_j\right)+1\right)e^{-\lambda_2\left(\left\lfloor \frac{\alpha_j}{T} \right\rfloor+1\right)T} - \\ &- e^{-\lambda_2(T+\beta_j)}\left(e^{\lambda_2 T}-1\right) - e^{-\lambda_2(T+\alpha_j)}\Bigg]e^{\lambda_2\left(\left\lfloor \frac{\alpha_j}{T} \right\rfloor+1\right)Tz}; \end{aligned}$$

$$A_{j,3}^{\text{uni}}(z) = z\sum_{k=0}^{\infty}\sum_{i=\left\lfloor\frac{\alpha_j}{T}\right\rfloor+2}^{\left\lfloor\frac{\beta_j}{T}\right\rfloor}\frac{\lambda_2 T - e^{\lambda_2(iT-\beta_j)}\left(1-e^{-\lambda_2 T}\right)}{\lambda_2\left(\beta_j-\alpha_j\right)}\cdot\frac{\left(\lambda_2 iTz\right)^k}{k!}e^{-\lambda_2 iT} =$$
$$= \frac{z}{\lambda_2\left(\beta_j-\alpha_j\right)}\left(e^{-\lambda_2\beta_j}\left(1-e^{-\lambda_2 T}\right)\frac{e^{\lambda_2\left(\left\lfloor\frac{\alpha_j}{T}\right\rfloor+1\right)Tz}-e^{\lambda_2\left\lfloor\frac{\beta_j}{T}\right\rfloor Tz}}{1-e^{-\lambda_2 Tz}}-\right.$$
$$\left.-\lambda_2 T\cdot\frac{e^{\lambda_2\left(\left\lfloor\frac{\alpha_j}{T}\right\rfloor+1\right)T(z-1)}-e^{\lambda_2\left\lfloor\frac{\beta_j}{T}\right\rfloor T(z-1)}}{1-e^{-\lambda_2 T(z-1)}}\right);$$

$$A_{j,4}^{\text{uni}}(z) = \frac{z}{\lambda_2\left(\beta_j-\alpha_j\right)}\left[\left(\lambda_2\beta_j-\lambda_2\left\lfloor\frac{\beta_j}{T}\right\rfloor T+1\right)e^{-\lambda_2\left(\left\lfloor\frac{\beta_j}{T}\right\rfloor+1\right)T}-\right.$$
$$\left.-e^{-\lambda_2(T+\beta_j)}\right]e^{\lambda_2\left(\left\lfloor\frac{\beta_j}{T}\right\rfloor+1\right)Tz}.$$

It should be noted that during the calculation of $A_{j,1}^{\text{uni}}(z)$ we have assumed that $\alpha_j \geq T$, but the final formula also holds if $\alpha_j < T$, as $A_{j,1}^{\text{uni}}(z) = 0$ in this case. A similar remark can be made regarding $A_{j,3}^{\text{uni}}(z)$: during calculation we have assumed that there are at least two multiples of T in the interval $(\alpha_j, \beta_j]$, but the final formula also holds if there is exactly one multiple of T in the interval (i. e. $\left\lfloor\frac{\alpha_j}{T}\right\rfloor + 1 = \left\lfloor\frac{\beta_j}{T}\right\rfloor$), as $A_{j,3}^{\text{uni}}(z) = 0$ in this case.

In the most general case, when there is at least one multiple of T in the interval $(\alpha_j, \beta_j]$, then

$$A_j^{\text{uni}}(z) = \pi_0^{\text{uni}} + A_{j,1}^{\text{uni}}(z) + A_{j,2}^{\text{uni}}(z) + A_{j,3}^{\text{uni}}(z) + A_{j,4}^{\text{uni}}(z)\,.$$

If there is no multiple of T in $(\alpha_j, \beta_j]$, then

$$A_j^{\text{uni}}(z) = \pi_0^{\text{uni}} + A_{j,1}^{\text{uni}}(z) + A_{j,5}^{\text{uni}}(z)\,,$$

where

$$A_{j,5}^{\text{uni}}(z) = z\left(e^{-\lambda_2\left(\left\lfloor\frac{\alpha_j}{T}\right\rfloor+1\right)T} - e^{-\lambda_2 T}\frac{e^{-\lambda_2\alpha_j}-e^{-\lambda_2\beta_j}}{\lambda_2\left(\beta_j-\alpha_j\right)}\right)e^{\lambda_2\left(\left\lfloor\frac{\alpha_j}{T}\right\rfloor+1\right)Tz}.$$

Generating function $B^{\text{uni}}(z)$ is going to be determined in a similar way to (2.8), but some extra terms have to be added. If u (the service time of the actual customer) falls in any interval $\big((i-1)\,T+y, iT+y\big]$, then the waiting time of the next customer is $iT+y$. Thus, the core basis of $E\big(z^{\xi}\mid y\big)$ is given by

$$E_0\big(z^{\xi}\mid y\big)=\sum_{k=0}^{\infty}\sum_{i=\left\lfloor\frac{\alpha_2}{T}\right\rfloor+1}^{\left\lfloor\frac{\beta_2}{T}\right\rfloor}\int\limits_{(i-1)T+y}^{iT+y}\frac{\left[\lambda_2\left(iT+y\right)z\right]^k}{k!}e^{-\lambda_2(iT+y)}\frac{\mathrm{d}u}{\beta_2-\alpha_2}=$$
$$=\frac{Te^{\lambda_2(z-1)y}}{\beta_2-\alpha_2}\cdot\frac{e^{\lambda_2(z-1)\left\lfloor\frac{\beta_2}{T}\right\rfloor T}-e^{\lambda_2(z-1)\left\lfloor\frac{\alpha_2}{T}\right\rfloor T}}{1-e^{-\lambda_2(z-1)T}},$$

and this is valid if there is at least one multiple of T in the interval $(\alpha_2,\beta_2]$. However, we have to make some adjustments to this around the boundaries of the interval. If $\alpha_2\le\left\lfloor\frac{\alpha_2}{T}\right\rfloor T+y$ (this happens with probability $\frac{e^{-\lambda_2\left(\alpha_2-\left\lfloor\frac{\alpha_2}{T}\right\rfloor T\right)}-e^{-\lambda_2 T}}{1-e^{-\lambda_2 T}}$), then an extra term has to be added to E_0; the waiting time in this case is $\left\lfloor\frac{\alpha_2}{T}\right\rfloor T+y$. If $\left\lfloor\frac{\alpha_2}{T}\right\rfloor T+y<\alpha_2$ (this happens with probability $\frac{1-e^{-\lambda_2\left(\alpha_2-\left\lfloor\frac{\alpha_2}{T}\right\rfloor T\right)}}{1-e^{-\lambda_2 T}}$), then an extra term has to be subtracted from E_0; the waiting time in this case is $\left(\left\lfloor\frac{\alpha_2}{T}\right\rfloor+1\right)T+y$. This additional term is

$$E_\alpha\big(z^{\xi}\mid y\big)=\sum_{k=0}^{\infty}\left[\int\limits_{\alpha_2}^{\left\lfloor\frac{\alpha_2}{T}\right\rfloor T+y}\frac{\left[\lambda_2\left(\left\lfloor\frac{\alpha_2}{T}\right\rfloor T+y\right)z\right]^k}{k!}e^{-\lambda_2\left(\left\lfloor\frac{\alpha_2}{T}\right\rfloor T+y\right)}\frac{\mathrm{d}u}{\beta_2-\alpha_2}\times\right.$$
$$\times\frac{e^{-\lambda_2\left(\alpha_2-\left\lfloor\frac{\alpha_2}{T}\right\rfloor T\right)}-e^{-\lambda_2 T}}{1-e^{-\lambda_2 T}}-$$
$$-\int\limits_{\left\lfloor\frac{\alpha_2}{T}\right\rfloor T+y}^{\alpha_2}\frac{\left[\lambda_2\left(\left(\left\lfloor\frac{\alpha_2}{T}\right\rfloor+1\right)T+y\right)z\right]^k}{k!}e^{-\lambda_2\left(\left(\left\lfloor\frac{\alpha_2}{T}\right\rfloor+1\right)T+y\right)}\frac{\mathrm{d}u}{\beta_2-\alpha_2}\times$$
$$\left.\times\frac{1-e^{-\lambda_2\left(\alpha_2-\left\lfloor\frac{\alpha_2}{T}\right\rfloor T\right)}}{1-e^{-\lambda_2 T}}\right]=$$
$$=\frac{\left\lfloor\frac{\alpha_2}{T}\right\rfloor T+y-\alpha_2}{\beta_2-\alpha_2}\cdot e^{\lambda_2(z-1)y}\times$$
$$\times\frac{e^{\lambda_2\left\lfloor\frac{\alpha_2}{T}\right\rfloor Tz-\lambda_2\alpha_2}\left(1-e^{\lambda_2(z-1)T}\right)-e^{\lambda_2\left\lfloor\frac{\alpha_2}{T}\right\rfloor T(z-1)-\lambda_2 T}\left(1-e^{\lambda_2 Tz}\right)}{1-e^{-\lambda_2 T}}.$$

The same method can be applied when adjusting the upper boundary:

$$E_\beta\left(z^\xi \mid y\right) = \sum_{k=0}^{\infty}\left[-\int\limits_{\beta_2}^{\lfloor\frac{\beta_2}{T}\rfloor T+y} \frac{\left[\lambda_2\left(\lfloor\frac{\beta_2}{T}\rfloor T+y\right)z\right]^k}{k!} e^{-\lambda_2\left(\lfloor\frac{\beta_2}{T}\rfloor T+y\right)}\frac{\mathrm{d}u}{\beta_2-\alpha_2}\times\right.$$

$$\times\frac{e^{-\lambda_2\left(\beta_2-\lfloor\frac{\beta_2}{T}\rfloor T\right)}-e^{-\lambda_2 T}}{1-e^{-\lambda_2 T}}+$$

$$+\int\limits_{\lfloor\frac{\beta_2}{T}\rfloor T+y}^{\beta_2} \frac{\left[\lambda_2\left(\left(\lfloor\frac{\beta_2}{T}\rfloor+1\right)T+y\right)z\right]^k}{k!} e^{-\lambda_2\left(\left(\lfloor\frac{\beta_2}{T}\rfloor+1\right)T+y\right)}\frac{\mathrm{d}u}{\beta_2-\alpha_2}\times$$

$$\left.\times\frac{1-e^{-\lambda_2\left(\beta_2-\lfloor\frac{\beta_2}{T}\rfloor T\right)}}{1-e^{-\lambda_2 T}}\right]=$$

$$=-\frac{\lfloor\frac{\beta_2}{T}\rfloor T+y-\beta_2}{\beta_2-\alpha_2}\cdot e^{\lambda_2(z-1)y}\times$$

$$\times\frac{e^{\lambda_2\lfloor\frac{\beta_2}{T}\rfloor Tz-\lambda_2\beta_2}\left(1-e^{\lambda_2(z-1)T}\right)-e^{\lambda_2\lfloor\frac{\beta_2}{T}\rfloor T(z-1)-\lambda_2 T}\left(1-e^{\lambda_2 Tz}\right)}{1-e^{-\lambda_2 T}}.$$

Therefore,

$$E\left(z^\xi \mid y\right)=E_\alpha\left(z^\xi \mid y\right)+E_0\left(z^\xi \mid y\right)+E_\beta\left(z^\xi \mid y\right).$$

Multiplying this expression by the density function of y and integrating with respect to y, we finally receive $B^{\mathrm{uni}}(z)$.

$$B_0^{\mathrm{uni}}(z)=\frac{T}{\beta_2-\alpha_2}\cdot\frac{e^{\lambda_2(z-2)T}-1}{\left(1-e^{-\lambda_2 T}\right)(z-2)}\cdot\frac{e^{\lambda_2\lfloor\frac{\beta_2}{T}\rfloor T(z-1)}-e^{\lambda_2\lfloor\frac{\alpha_2}{T}\rfloor T(z-1)}}{1-e^{-\lambda_2(z-1)T}};$$

$$B_\alpha^{\mathrm{uni}}(z)=\frac{e^{\lambda_2\lfloor\frac{\alpha_2}{T}\rfloor Tz-\lambda_2\alpha_2}\left(1-e^{\lambda_2(z-1)T}\right)-e^{\lambda_2\lfloor\frac{\alpha_2}{T}\rfloor T(z-1)-\lambda_2 T}\left(1-e^{\lambda_2 Tz}\right)}{\left(\beta_2-\alpha_2\right)\left(1-e^{-\lambda_2 T}\right)^2}\times$$
$$\times\left(\frac{\left(\lfloor\frac{\alpha_2}{T}\rfloor T-\alpha_2\right)\left(e^{\lambda_2(z-2)T}-1\right)+Te^{\lambda_2(z-2)T}}{z-2}+\frac{1-e^{\lambda_2(z-2)T}}{\lambda_2(z-2)^2}\right);$$

$$B_\beta^{\mathrm{uni}}(z)=-\frac{e^{\lambda_2\lfloor\frac{\beta_2}{T}\rfloor Tz-\lambda_2\beta_2}\left(1-e^{\lambda_2(z-1)T}\right)-e^{\lambda_2\lfloor\frac{\beta_2}{T}\rfloor T(z-1)-\lambda_2 T}\left(1-e^{\lambda_2 Tz}\right)}{\left(\beta_2-\alpha_2\right)\left(1-e^{-\lambda_2 T}\right)^2}\times$$
$$\times\left(\frac{\left(\lfloor\frac{\beta_2}{T}\rfloor T-\beta_2\right)\left(e^{\lambda_2(z-2)T}-1\right)+Te^{\lambda_2(z-2)T}}{z-2}+\frac{1-e^{\lambda_2(z-2)T}}{\lambda_2(z-2)^2}\right).$$

With these

$$B^{\text{uni}}(z) = B_{\alpha}^{\text{uni}}(z) + B_{0}^{\text{uni}}(z) + B_{\beta}^{\text{uni}}(z)\,.$$

We still must not forget about the opportunity when there is no multiple of T in the interval $(\alpha_2, \beta_2]$. In this case $\left\lfloor \frac{\alpha_2}{T} \right\rfloor = \left\lfloor \frac{\beta_2}{T} \right\rfloor$, and three subcases will be considered depending on where y in the interval $[0, T)$ is situated relative to α_2 and β_2 mod T.

If $y < \alpha_2 - \left\lfloor \frac{\alpha_2}{T} \right\rfloor T$ (this happens with probability $\frac{1-e^{-\lambda_2\left(\alpha_2 - \left\lfloor \frac{\alpha_2}{T} \right\rfloor T\right)}}{1-e^{-\lambda_2 T}}$), then the time elapsed between two consecutive starts of services is $\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor + 1\right) T + y$:

$$E_1\left(z^{\xi} \mid y\right) = \sum_{k=0}^{\infty} \int\limits_{\alpha_2}^{\beta_2} \frac{\left[\lambda_2\left(\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor + 1\right) T + y\right) z\right]^k}{k!} e^{-\lambda_2\left(\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor + 1\right) T + y\right)} \frac{\mathrm{d}u}{\beta_2 - \alpha_2} \times$$
$$\times \frac{1 - e^{-\lambda_2\left(\alpha_2 - \left\lfloor \frac{\alpha_2}{T} \right\rfloor T\right)}}{1 - e^{-\lambda_2 T}} = e^{\lambda_2\left(\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor + 1\right) T + y\right)(z-1)} \cdot \frac{1 - e^{-\lambda_2\left(\alpha_2 - \left\lfloor \frac{\alpha_2}{T} \right\rfloor T\right)}}{1 - e^{-\lambda_2 T}}.$$

If $\alpha_2 - \left\lfloor \frac{\alpha_2}{T} \right\rfloor T \le y < \beta_2 - \left\lfloor \frac{\beta_2}{T} \right\rfloor T$, then interval $[\alpha_2, \beta_2]$ is divided into two parts by $\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y$; it is obvious that the waiting time is one cycle longer with u being in one of them, than in the other:

$$E_2\left(z^{\xi} \mid y\right) = \sum_{k=0}^{\infty} \left[\int\limits_{\alpha_2}^{\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y} \frac{\left[\lambda_2\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y\right) z\right]^k}{k!} e^{-\lambda_2\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y\right)} \frac{\mathrm{d}u}{\beta_2 - \alpha_2} + \right.$$
$$\left. \int\limits_{\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y}^{\beta_2} \frac{\left[\lambda_2\left(\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor + 1\right) T + y\right) z\right]^k}{k!} e^{-\lambda_2\left(\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor + 1\right) T + y\right)} \frac{\mathrm{d}u}{\beta_2 - \alpha_2} \right] \times$$
$$\times \frac{e^{-\lambda_2\left(\alpha_2 - \left\lfloor \frac{\alpha_2}{T} \right\rfloor T\right)} - e^{-\lambda_2\left(\beta_2 - \left\lfloor \frac{\beta_2}{T} \right\rfloor T\right)}}{1 - e^{-\lambda_2 T}} =$$
$$= \frac{\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y - \alpha_2\right) + \left(\beta_2 - \left\lfloor \frac{\alpha_2}{T} \right\rfloor T - y\right) e^{\lambda_2 (z-1) T}}{\beta_2 - \alpha_2} e^{\lambda_2\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y\right)(z-1)} \times$$
$$\times \frac{e^{-\lambda_2\left(\alpha_2 - \left\lfloor \frac{\alpha_2}{T} \right\rfloor T\right)} - e^{-\lambda_2\left(\beta_2 - \left\lfloor \frac{\beta_2}{T} \right\rfloor T\right)}}{1 - e^{-\lambda_2 T}}.$$

And finally, if $y \geq \beta_2 - \left\lfloor \frac{\beta_2}{T} \right\rfloor T$, then:

$$
\begin{aligned}
E_3\big(z^\xi \mid y\big) &= \sum_{k=0}^{\infty} \int_{\alpha_2}^{\beta_2} \frac{\left[\lambda_2 \left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y\right) z\right]^k}{k!} e^{-\lambda_2\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y\right)} \frac{\mathrm{d}u}{\beta_2 - \alpha_2} \times \\
&\times \frac{e^{-\lambda_2\left(\beta_2 - \left\lfloor \frac{\beta_2}{T} \right\rfloor T\right)} - e^{-\lambda_2 T}}{1 - e^{-\lambda_2 T}} = e^{\lambda_2\left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y\right)(z-1)} \cdot \frac{e^{-\lambda_2\left(\beta_2 - \left\lfloor \frac{\beta_2}{T} \right\rfloor T\right)} - e^{-\lambda_2 T}}{1 - e^{-\lambda_2 T}}.
\end{aligned}
$$

Using these notations

$$
\begin{aligned}
E\big(z^\xi \mid y\big) &= E_1\big(z^\xi \mid y\big) + E_2\big(z^\xi \mid y\big) + E_3\big(z^\xi \mid y\big) = \\
&= \frac{e^{\lambda_2 \left\lfloor \frac{\alpha_2}{T} \right\rfloor T z} e^{\lambda_2 (z-1) y}}{1 - e^{-\lambda_2 T}} \Bigg[e^{\lambda_2 (z-1) T} \left(e^{-\lambda_2 \left\lfloor \frac{\alpha_2}{T} \right\rfloor T} - e^{-\lambda_2 \alpha_2} \right) + \\
&+ \left(-\alpha_2 + \beta_2 e^{\lambda_2 (z-1) T} + \left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor T + y \right) \left(1 - e^{\lambda_2 (z-1) T} \right) \right) \frac{e^{-\lambda_2 \alpha_2} - e^{-\lambda_2 \beta_2}}{\beta_2 - \alpha_2} + \\
&\qquad + e^{-\lambda_2 \beta_2} - e^{-\lambda_2 \left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor + 1 \right) T} \Bigg].
\end{aligned}
$$

The generating function in this case is

$$
\begin{aligned}
B^{\mathrm{uni}}(z) &= \frac{e^{\lambda_2 \left\lfloor \frac{\alpha_2}{T} \right\rfloor T z}}{\left(1 - e^{-\lambda_2 T}\right)^2} \frac{e^{\lambda_2 (z-2) T} - 1}{z - 2} \Bigg[e^{\lambda_2 (z-1) T} \left(e^{-\lambda_2 \left\lfloor \frac{\alpha_2}{T} \right\rfloor T} - e^{-\lambda_2 \alpha_2} \right) + \\
&+ \left(-\alpha_2 + \beta_2 e^{\lambda_2 (z-1) T} \right) \frac{e^{-\lambda_2 \alpha_2} - e^{-\lambda_2 \beta_2}}{\beta_2 - \alpha_2} + e^{-\lambda_2 \beta_2} - e^{-\lambda_2 \left(\left\lfloor \frac{\alpha_2}{T} \right\rfloor + 1 \right) T} \Bigg] + \\
&+ \frac{e^{\lambda_2 \left\lfloor \frac{\alpha_2}{T} \right\rfloor T z}}{\left(1 - e^{-\lambda_2 T}\right)^2} \left(e^{\lambda_2 (z-1) T} - 1 \right) \frac{e^{-\lambda_2 \alpha_2} - e^{-\lambda_2 \beta_2}}{\beta_2 - \alpha_2} \times \\
&\times \left(\frac{\left\lfloor \frac{\alpha_2}{T} \right\rfloor T \left(e^{\lambda_2 (z-2) T} - 1 \right) + T e^{\lambda_2 (z-2) T}}{z - 2} + \frac{1 - e^{\lambda_2 (z-2) T}}{\lambda_2 (z-2)^2} \right).
\end{aligned}
$$

It is obvious, that generating functions (2.6) of $C(z)$ and (2.9) of $P(z)$ are also valid with newly given generating functions $A_j^{\mathrm{uni}}(z)$ and $B^{\mathrm{uni}}(z)$.

2.3 Validating results

Equilibrium probabilities of states 0 and 1

In (2.10) the equilibrium probability of free state was determined as the function of the derivatives of the generating functions while $z \to 1$, which were given in (2.11). Probability p_1 was given in (2.13), whereas other equilibrium probabilities can be derived from (2.9).

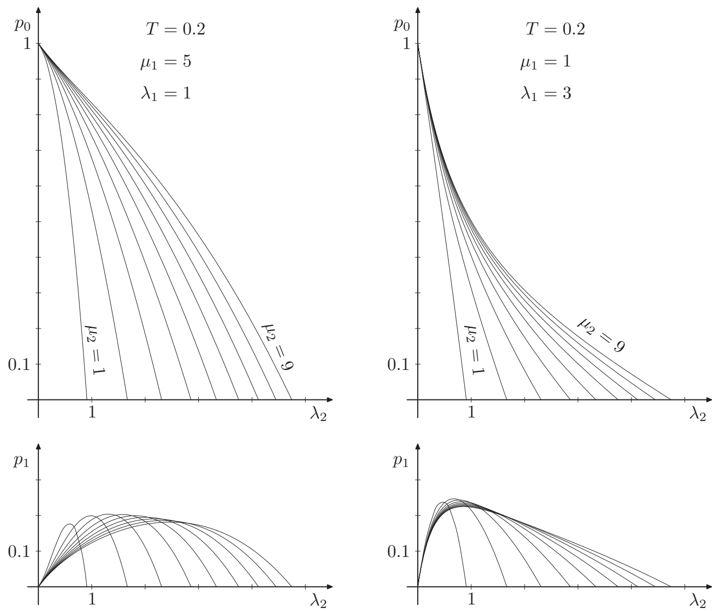


Figure 2.4: *Probabilities p_0 and p_1 with exponentially distributed service times*

It is advisable to have a closer look at the first two, explicitly given probabilities, with special attention to how they vary with the change of the input parameters. Considering the high number of input parameters and our extra assumption on the boundaries of the uniform distributions, this seems feasible only with the service time distribution of second-type customers being exponential. Figures 2.4 and 2.5 show their behaviour when the service time of first-type customers is exponentially and uniformly distributed, respectively. The load intensity of second-type customers is chosen as the main independent variable at nine different values of their service capacity ($\mu_2 = 1, 2, \ldots, 9$). Two sets of parameters of first-type customers are chosen to have the opportunity to compare the effect of a weak load on a high-capacity system, and a heavy load on a low-capacity one.

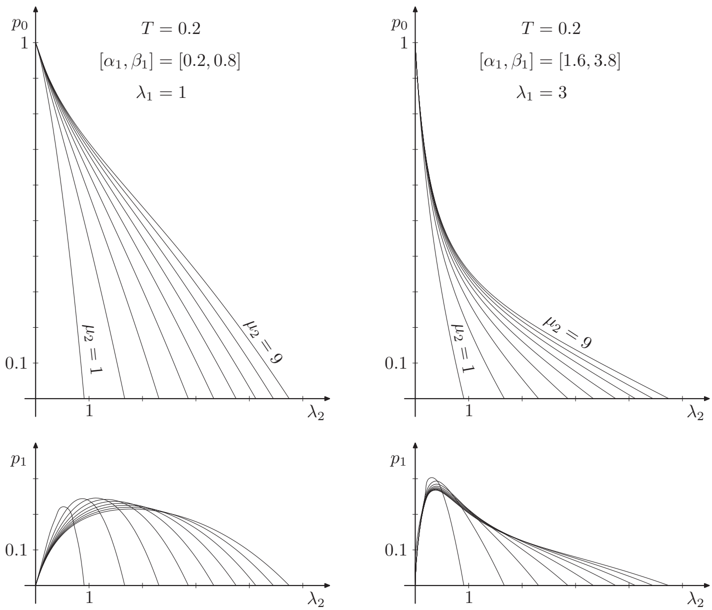


Figure 2.5: *Probabilities p_0 and p_1 with uniformly distributed first-type service time*

As it is expected, several relevant observations can be made. Obviously, the probability of free state is a decreasing function of the input intensity; whereas p_1 has a maximum between 0 and $\lambda_2^{\max}$, where $\lambda_2^{\max}$ is the largest possible value of the input intensity allowed by ergodicity condition (2.14). With chosen parameters ($T = 0.2$ and $\mu_2 = 1, 2, \ldots, 9$) these are 0.9, 1.7, 2.3, 2.8, 3.3, 3.7, 4.1, 4.4, 4.7 (truncated to the first significant decimal), respectively. Values of $\lambda_2^{\max}$ for other parameters have been seen previously in Figure 2.3.

Not surprisingly, although first-type customers do have a considerable effect on the first two equilibrium probabilities (they decrease at lower values of λ_2 if first-type load is heavier), but they have no effect on the ergodicity of the process; all functions take their zero-points at the same values. This is in accordance with (2.14), as it contains no first-type parameters.

Numerical investigation

Computer experiments were carried out with different values of parameters of arrival rate and service time. For all fixed values of the parameters 1000 independent experiments were made with randomly generated arrival and service times. Within the experiment we determined the number of requests waiting for service at each instant when the service of a particular request started. Data were generated in such a way, that we were able to examine the behaviour of the system serving at least 120 entities in each of the 1000 experiments.

Figures show results in the case of a server serving two types of customers with exponentially or uniformly distributed service times. For the two types of customers λ_j denotes arrival rate, service time distributions are either exponential with parameters μ_j or uniform in the intervals $[\alpha_j, \beta_j]$, and T is cycle-time ($j = 1, 2$).

Graphs display only the first few values of p_i, as other values of p_i are very small, and would require a much larger number of experiments. The graph-plot figures also contain theoretically calculated values of p_i, they are represented as straight lines. It can be seen that numerical results are in high accordance with theoretical values.

Maple 7 was used to carry out the numerical investigation. Uniform service time and exponential inter-arrival and service time distributions were provided through the built-in methods of the program. The exponential variate is generated from the uniform variate based on the relation

$$\text{Exponential}\,(\lambda) \sim -\frac{1}{\lambda}\ln\left(\text{Uniform}\,(0,1)\right).$$

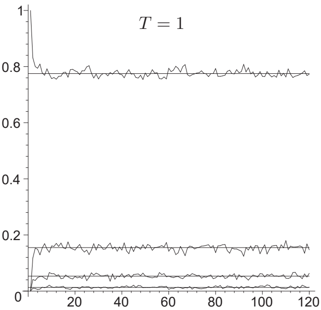
Experiments were carried out on a personal computer with a CPU of 3.06 GHz and a RAM of 504 MB; on average 20–25 minutes were necessary to complete the experiment with each setting.

The results of four numerical experiments are presented for each case.

Case 1: Both service time distributions are exponential.

$\lambda_1 = 0.4 \quad \mu_1 = 1$

$\lambda_2 = 0.2 \quad \mu_2 = 1$

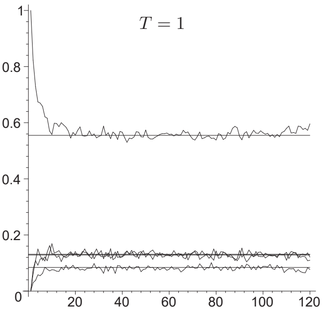


$p_0 = 0.775 \quad p_1 = 0.155$

$p_2 = 0.053 \quad p_3 = 0.013$

$\lambda_1 = 0.5 \quad \mu_1 = 5$

$\lambda_2 = 0.8 \quad \mu_2 = 3$

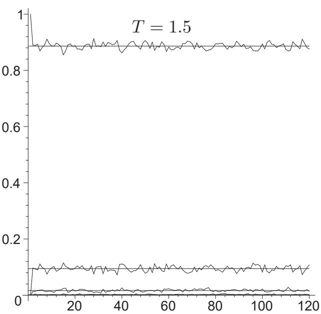


$p_0 = 0.555 \quad p_1 = 0.128$

$p_2 = 0.131 \quad p_3 = 0.084$

$\lambda_1 = 0.4 \quad \mu_1 = 1$

$\lambda_2 = 0.1 \quad \mu_2 = 0.8$

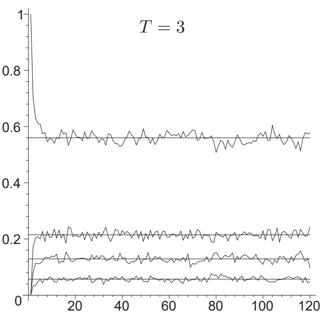


$p_0 = 0.886 \quad p_1 = 0.094$

$p_2 = 0.017 \quad p_3 = 0.002$

$\lambda_1 = 0.5 \quad \mu_1 = 0.6$

$\lambda_2 = 0.3 \quad \mu_2 = 1$



$p_0 = 0.560 \quad p_1 = 0.214$

$p_2 = 0.129 \quad p_3 = 0.057$

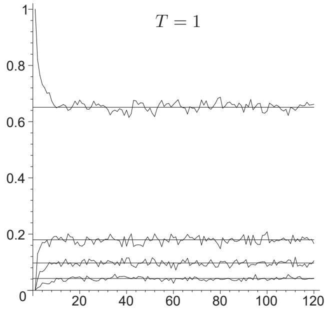
Figure 2.6: *Numerical results; case 1*

Case 2: Service time distribution of first-type customers is exponential, and that of second-type ones is uniform.

$\lambda_1 = 0.4 \quad \mu_1 = 3$

$\lambda_2 = 0.2 \quad \alpha_2 = 1 \quad \beta_2 = 4$

$T = 1$



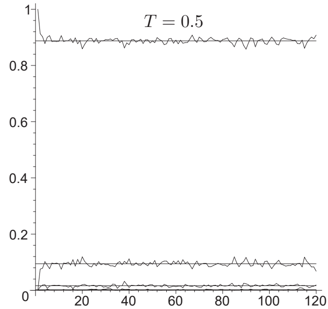
$p_0 = 0.651 \quad p_1 = 0.180$

$p_2 = 0.097 \quad p_3 = 0.042$

$\lambda_1 = 0.6 \quad \mu_1 = 1.3$

$\lambda_2 = 0.1 \quad \alpha_2 = 1.5 \quad \beta_2 = 2.5$

$T = 0.5$



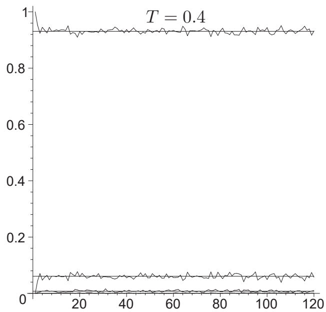
$p_0 = 0.887 \quad p_1 = 0.094$

$p_2 = 0.016 \quad p_3 = 0.002$

$\lambda_1 = 0.3 \quad \mu_1 = 2.5$

$\lambda_2 = 0.08 \quad \alpha_2 = 1.6 \quad \beta_2 = 2.4$

$T = 0.4$



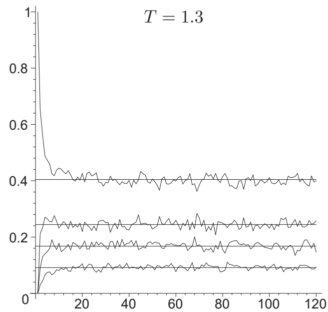
$p_0 = 0.931 \quad p_1 = 0.061$

$p_2 = 0.008 \quad p_3 = 0.001$

$\lambda_1 = 0.6 \quad \mu_1 = 0.4$

$\lambda_2 = 0.2 \quad \alpha_2 = 1.3 \quad \beta_2 = 4.2$

$T = 1.3$



$p_0 = 0.404 \quad p_1 = 0.244$

$p_2 = 0.168 \quad p_3 = 0.092$

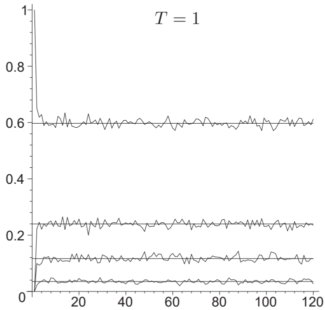
Figure 2.7: *Numerical results; case 2*

Case 3: Service time distribution of first-type customers is uniform, and that of second-type ones is exponential.

$\lambda_1 = 0.7 \quad \alpha_1 = 1 \quad \beta_1 = 3$

$\lambda_2 = 0.3 \quad \mu_2 = 2$

$T = 1$



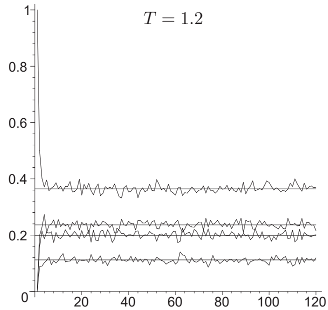
$p_0 = 0.597 \quad p_1 = 0.240$

$p_2 = 0.117 \quad p_3 = 0.035$

$\lambda_1 = 0.5 \quad \alpha_1 = 2.4 \quad \beta_1 = 4.8$

$\lambda_2 = 0.4 \quad \mu_2 = 1.5$

$T = 1.2$



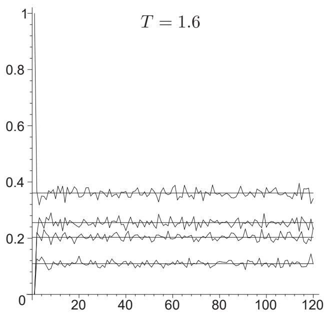
$p_0 = 0.364 \quad p_1 = 0.236$

$p_2 = 0.199 \quad p_3 = 0.112$

$\lambda_1 = 0.8 \quad \alpha_1 = 4.8 \quad \beta_1 = 6.4$

$\lambda_2 = 0.3 \quad \mu_2 = 2.2$

$T = 1.6$



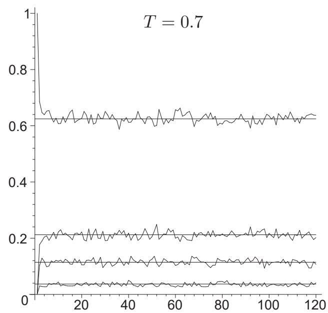
$p_0 = 0.361 \quad p_1 = 0.255$

$p_2 = 0.203 \quad p_3 = 0.109$

$\lambda_1 = 0.6 \quad \alpha_1 = 0.7 \quad \beta_1 = 2.1$

$\lambda_2 = 0.4 \quad \mu_2 = 3$

$T = 0.7$



$p_0 = 0.624 \quad p_1 = 0.212$

$p_2 = 0.114 \quad p_3 = 0.037$

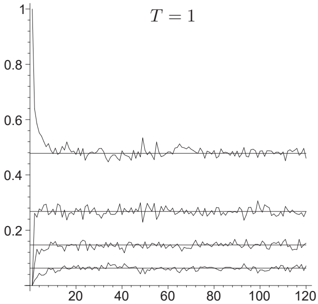
Figure 2.8: *Numerical results; case 3*

Case 4: Both service time distributions are uniform.

$\lambda_1 = 0.4 \quad \alpha_1 = 1 \quad \beta_1 = 3$

$\lambda_2 = 0.2 \quad \alpha_2 = 1 \quad \beta_2 = 4$

$T = 1$



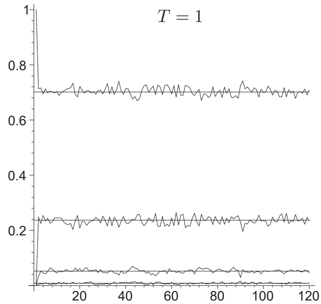
$p_0 = 0.478 \quad p_1 = 0.268$

$p_2 = 0.147 \quad p_3 = 0.063$

$\lambda_1 = 0.5 \quad \alpha_1 = 5 \quad \beta_1 = 7$

$\lambda_2 = 0.05 \quad \alpha_2 = 3 \quad \beta_2 = 8$

$T = 1$



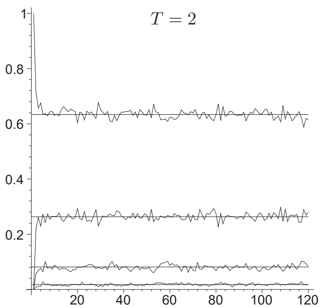
$p_0 = 0.701 \quad p_1 = 0.237$

$p_2 = 0.052 \quad p_3 = 0.008$

$\lambda_1 = 0.3 \quad \alpha_1 = 2 \quad \beta_1 = 6$

$\lambda_2 = 0.1 \quad \alpha_2 = 2 \quad \beta_2 = 4$

$T = 2$



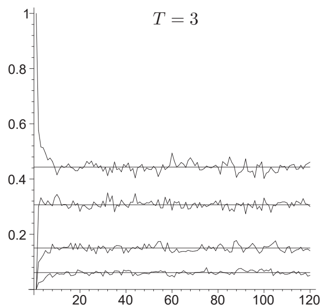
$p_0 = 0.634 \quad p_1 = 0.263$

$p_2 = 0.081 \quad p_3 = 0.018$

$\lambda_1 = 0.2 \quad \alpha_1 = 6 \quad \beta_1 = 12$

$\lambda_2 = 0.06 \quad \alpha_2 = 3 \quad \beta_2 = 15$

$T = 3$



$p_0 = 0.443 \quad p_1 = 0.307$

$p_2 = 0.150 \quad p_3 = 0.061$

Figure 2.9: *Numerical results; case 4*

Chapter 3

Discrete cycling-waiting retrial systems

In this chapter a similar Lakatos-type system is to be investigated as in Chapter 2, but inter-arrival and service time distributions are not continuous, but discrete. In the light of technical applications it is even more desirable to extend the description of such systems for discrete distributions.

Consider a Lakatos-type queuing system serving two types of customers. Cycle-time T is divided into n equal time-slices. The probability of appearance of a j^{th}-type customer during a certain time-slice is r_j, so there is no entry with probability $1 - r_j$, i. e. inter-arrival times are geometrically distributed with parameters r_j $(j = 1, 2)$. There are several aspects which can be treated just exactly in the same way as in the case of continuous distributions. To describe the system, we use the same embedded Markov-chain (the number of customers in the system just before the service of a customer begins); and to see that the process is still Markovian we refer to the same argument and the memoryless property of the geometric distribution. The same transition probabilities and their generating functions are introduced, thus, the matrix of transition probabilities is also (2.1).

However, there are some differences between discrete and continuous systems, out of which a significant phenomenon is collision. With continuously distributed service times the probability of appearance of two customers of different types at the very same instant is 0, but in discrete systems different-type customers do appear during the same time-slice with non-zero probability. There are several ways to deal with collisions; we opt for three methods treating them, numbered I., II., and III., and we will refer to them by these numbers from now on.

Collision treatment method I. In case of collision both types of customers are refused.

Collision treatment method II. In case of collision first-type customers are accepted for service, but second-type ones are refused.

Collision treatment method III. In case of collision customers of both types are accepted for service, but first-type ones are served first. When applying this method, in addition to previously defined transition probabilities new ones have to be introduced. Let a_{12i} denote the probability of appearance of i customers of second-type at the service of a first-type customer, if the service process started with the simultaneous appearance of customers of both types; the generating function of these probabilities is $A_{12}(z) = \sum_{i=0}^{\infty} a_{12i} z^i$.

The aim is to determine generating functions of transition probabilities, as well as to establish condition of ergodicity. Two service time distributions are considered; for one possibility service times are geometrically distributed with parameters q_j, i. e. the service of a j^{th}-type customer continues during a time unit with probability q_j, and terminates with probability $1 - q_j$. The other examined alternative is when service time distributions are uniform in the intervals $[\gamma_j, \delta_j]$, where γ_j and δ_j are multiples of T, i. e. the probability that the service of a j^{th}-type customer is any time units in this interval is $q_j = \frac{T}{n(\delta_j - \gamma_j)}$. All three collision disciplines are investigated with each system.

3.1 Theoretical results

Consider a discrete cyclic-waiting system serving two types of customers in which inter-arrival time distributions are geometric with parameters r_j ($j =$

$1,2)$. Service time distributions of customers of either type may be geometric with parameters q_i or uniform in the intervals $[\gamma_j, \delta_j]$; thus (considering the three collision disciplines) 12 different cases are to be examined $(i, j = 1, 2)$. The service of an entering customer may start immediately on arrival if the server is free, but in case of a busy server or waiting customers, first-type customers are refused, and second-type customers join the queue. The service of queued customers may start at times differing from their arrival times by multiples of cycle-time T, which is divided into n equal time slices; these form the units of the probability distributions. The FIFO (FCFS) rule is obeyed. The states of the corresponding embedded Markov-chain are identified with the number of customers in the system at moments just before starting the service of a customer.

Theorem 3.1.1 *The matrix of transition probabilities of this chain has the same form* (2.1).

The generating functions of the defined transition probabilities are given below. The type of service time distribution is indicated in the upper index: $A_j^{\{geo,uni\}}(z)$ *indicates the type of service time distribution of* j^{th}*-type customers,* $A_{12}^{\{geo,uni\}}(z)$ *indicates the type of service time distribution of first-type customers, and* $B^{\{geo,uni\}}(z)$ *indicates the type of service time distribution of second-type customers.*

$$A_j^{geo}(z) = \frac{(1-r_2)(1-q_j)}{1-q_j(1-r_2)} + z\frac{r_2(1-q_j)}{1-q_j(1-r_2)} + \\ + z\frac{r_2 q_j \left(1-q_j^n\right)(1-r_2+r_2 z)^n}{(1-q_j(1-r_2))\left(1-q_j^n(1-r_2+r_2 z)^n\right)}, \tag{3.1}$$

$$A_j^{uni}(z) = \frac{q_j}{r_2}\left[(1-r_2)^{\frac{\gamma_j}{T}n+1} - (1-r_2)^{\frac{\delta_j}{T}n+1}\right] + \\ + zq_j\left[(1-r_2)^{\frac{\gamma_j}{T}n} - (1-r_2)^{\frac{\delta_j}{T}n}\right] + zq_j\left[(1-r_2)^{\frac{\gamma_j}{T}n} - (1-r_2)^{\frac{\delta_j}{T}n}\right] \times \\ \times \frac{1-(1-r_2)^n}{r_2(1-r_2)^n}(1-r_2+r_2 z)^n \frac{1-\left(\frac{1-r_2+r_2 z}{1-r_2}\right)^{\frac{\gamma_j}{T}n}}{1-\left(\frac{1-r_2+r_2 z}{1-r_2}\right)^n} +$$

$$+ zq_j\left(1-r_2+r_2z\right)^n\left[n\frac{\left(1-r_2+r_2z\right)^{\frac{\gamma_j}{T}n}-\left(1-r_2+r_2z\right)^{\frac{\delta_j}{T}n}}{1-\left(1-r_2+r_2z\right)^n}-\right.$$

$$\left.-\left(1-r_2\right)^{\frac{\delta_j}{T}n}\frac{1-\left(1-r_2\right)^n}{r_2\left(1-r_2\right)^n}\frac{\left(\frac{1-r_2+r_2z}{1-r_2}\right)^{\frac{\gamma_j}{T}n}-\left(\frac{1-r_2+r_2z}{1-r_2}\right)^{\frac{\delta_j}{T}n}}{1-\left(\frac{1-r_2+r_2z}{1-r_2}\right)^n}\right], \quad (3.2)$$

$$A_{12}^{geo}(z) = z\frac{\left(1-q_1^n\right)\left(1-r_2+r_2z\right)^n}{1-q_1^n\left(1-r_2+r_2z\right)^n}, \quad (3.3)$$

$$A_{12}^{uni}(z) = znq_1\frac{\left(1-r_2+r_2z\right)^{\left(\frac{\gamma_1}{T}+1\right)n}-\left(1-r_2+r_2z\right)^{\left(\frac{\delta_1}{T}+1\right)n}}{1-\left(1-r_2+r_2z\right)^n}, \quad (3.4)$$

$$B^{geo}(z) = \frac{1-\left(1-r_2\right)^n\left(1-r_2+r_2z\right)^n}{1-\left(1-r_2\right)\left(1-r_2+r_2z\right)}\cdot\frac{r_2\left(1-r_2+r_2z\right)}{1-\left(1-r_2\right)^n}+$$
$$+\frac{1-q_2^n\left(1-r_2\right)^n\left(1-r_2+r_2z\right)^n}{1-q_2\left(1-r_2\right)\left(1-r_2+r_2z\right)}\times$$
$$\times\frac{r_2q_2\left(1-r_2+r_2z\right)\left(\left(1-r_2+r_2z\right)^n-1\right)}{\left(1-\left(1-r_2\right)^n\right)\left(1-\left(1-r_2\right)^n\left(1-r_2+r_2z\right)^n\right)}, \quad (3.5)$$

$$B^{uni}(z) = \frac{r_2q_2}{1-\left(1-r_2^n\right)}\frac{\left(1-r_2+r_2z\right)^{\frac{\gamma_2}{T}n}-\left(1-r_2+r_2z\right)^{\frac{\delta_2}{T}n}}{1-\left(1-r_2+r_2z\right)^n}\times$$
$$\times\left[\left(\left(1-r_2+r_2z\right)-\left(1-r_2+r_2z\right)^{n+1}\right)\times\right.$$
$$\times\left(\frac{1-\left(1-r_2\right)^n\left(1-r_2+r_2z\right)^n}{\left(1-\left(1-r_2\right)\left(1-r_2+r_2z\right)\right)^2}--\frac{n\left(1-r_2\right)^n\left(1-r_2+r_2z\right)^n}{1-\left(1-r_2\right)\left(1-r_2+r_2z\right)}\right)+$$
$$\left.+n\left(1-r_2+r_2z\right)^{n+1}\frac{1-\left(1-r_2\right)^n\left(1-r_2+r_2z\right)^n}{1-\left(1-r_2\right)\left(1-r_2+r_2z\right)}\right] \quad (3.6)$$

and $C(z)$ depends on collision policies:

I. $C(z) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2}A_1(z)+\frac{r_2(1-r_1)}{r_1+r_2-r_1r_2}A_2(z)+\frac{r_1r_2}{r_1+r_2-r_1r_2}$, (3.7a)

II. $C(z) = \frac{r_1}{r_1+r_2-r_1r_2}A_1(z)+\frac{r_2(1-r_1)}{r_1+r_2-r_1r_2}A_2(z)$, (3.7b)

III. $C(z) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2}A_1(z)+\frac{r_2(1-r_1)}{r_1+r_2-r_1r_2}A_2(z)+\frac{r_1r_2}{r_1+r_2-r_1r_2}A_{12}(z)$. (3.7c)

Proof A similar routine is followed as in the continuous case, but peculiarities of discrete systems are born in mind.

First we determine a_{ji}. In this case only one customer is present at the beginning of a service (the one whose service is about to start). Time units are of length $\frac{T}{n}$, and all time intervals are measured by them. The service of the actual customer is denoted by u, and the next one appears v times after its service started. In order to get a_{ji}, the distribution of $u-v$ must be known. In the case of geometric distribution

$$P^{\mathrm{geo}}(u-v=l)=\sum_{k=l+1}^{\infty} q_j^{k-1}\left(1-q_j\right)\left(1-r_2\right)^{k-l-1} r_2=\frac{r_2\left(1-q_j\right) q_j^l}{1-q_j\left(1-r_2\right)}.$$

If service time is uniformly distributed, then two separate calculations have to be carried out.
If $0<l \leq \frac{\gamma_j}{T} n$:

$$P^{\mathrm{uni}}(u-v=l)=\sum_{k=\frac{\gamma_j}{T} n+1}^{\frac{\delta_j}{T}} q_j\left(1-r_2\right)^{k-l-1} r_2= \\ =q_j\left[(1-r)^{\frac{\gamma_j}{T} n-l}-(1-r)^{\frac{\delta_j}{T} n-l}\right];$$

and if $\frac{\gamma_j}{T} n<l \leq \frac{\delta_j}{T} n$:

$$P^{\mathrm{uni}}(u-v=l)=\sum_{k=l+1}^{\frac{\delta_j}{T}} q_j\left(1-r_2\right)^{k-l-1} r_2=q_j\left[1-(1-r)^{\frac{\delta_j}{T} n-l}\right].$$

The waiting time can be determined on the basis of these probabilities. If $u-v=0$ (the next customer appears in the time-slice in which the service of the present customer is completed), then it can start its service immediately, the waiting time is 0. If $u-v$ is in $\overline{1, n}$, then the waiting time is T, i. e. n units; if it is in $\overline{n+1,2n}$, then the waiting time is $2n$; and in general if $u-v$ takes on some value from $\overline{(i-1) n+1, in}$, then the waiting time of the next customer is in.

Take the case of geometric distribution first; the corresponding probabilities are

$$\frac{r_2\left(1-q_j\right)}{1-q_j\left(1-r_2\right)}, \quad \frac{r_2\left(1-q_j^n\right)}{1-q_j\left(1-r_2\right)}, \quad \frac{r_2\left(q_j^n-q_j^{2n}\right)}{1-q_j\left(1-r_2\right)}, \quad \ldots,$$

and in general

$$P^{\text{geo}}((i-1)\,n+1 \le u-v \le in) = \sum_{l=(i-1)n+1}^{in} \frac{r_2\,(1-q_j)\,q_j^l}{1-q_j\,(1-r_2)} =$$

$$= \frac{r_2\,(1-q_j)}{1-q_j\,(1-r_2)} q_j^{(i-1)n+1} \frac{1-q_j^n}{1-q_j} = \frac{r_2 q_j\left(1-q_j^n\right)}{1-q_j\,(1-r_2)} q_j^{(i-1)n}.$$

The generating function of number of customers appearing in a time slice is $1-r_2+r_2z$, hence, the generating function of number of customers entering during the waiting time is

$$\sum_{i=1}^{\infty} \frac{r_2 q_j\left(1-q_j^n\right)}{1-q_j\,(1-r_2)} q_j^{(i-1)n}\,(1-r_2+r_2z)^{in} =$$

$$= \frac{r_2 q_j\,(1-r_2+r_2z)^n\left(1-q_j^n\right)}{(1-q_j\,(1-r_2))\left(1-q_j^n\,(1-r_2+r_2z)^n\right)}.$$

Taking into account that we examined only those events when at least one customer enters, this expression has to be multiplied by z. Considering also that the probability that the waiting time is zero (the next customer enters in the last time-slice, the system stays in state 1) is $\frac{r_2(1-q_j)}{1-q_j(1-r_2)}$, as well as that there is no entry at all (the system returns to free state) with probability $\frac{(1-r_2)(1-q_j)}{1-q_j(1-r_2)}$, we get (3.1).

In the case of uniform distribution the probability that the waiting time of the second customer is in (i. e. $u-v$ is in $\overline{(i-1)\,n+1, in)}$ is the following. If $0 < i \le \frac{\gamma_j}{T}$:

$$P^{\text{uni}}\big(\,(i-1)\,n+1 \le u-v \le in\big) =$$

$$= \sum_{l=(i-1)n+1}^{in} q_j \left[(1-r_2)^{\frac{\gamma_j}{T}n-l} - (1-r_2)^{\frac{\delta_j}{T}n-l}\right] =$$

$$= q_j \left[(1-r_2)^{\frac{\gamma_j}{T}n} - (1-r_2)^{\frac{\delta_j}{T}n}\right] \frac{1-(1-r_2)^n}{r_2\,(1-r_2)^n} \frac{1}{(1-r_2)^{(i-1)n}},$$

thus, the generating function of entering customers is

$$\sum_{i=1}^{\frac{\gamma_j}{T}} q_j \left[(1-r_2)^{\frac{\gamma_j}{T}n} - (1-r_2)^{\frac{\delta_j}{T}n} \right] \frac{1-(1-r_2)^n}{r_2(1-r_2)^n} \frac{(1-r_2+r_2z)^{in}}{(1-r_2)^{(i-1)n}} =$$

$$= q_j \left[(1-r_2)^{\frac{\gamma_j}{T}n} - (1-r_2)^{\frac{\delta_j}{T}n} \right] \times$$

$$\times \frac{1-(1-r_2)^n}{r_2(1-r_2)^n} (1-r_2+r_2z)^n \frac{1-\left(\frac{1-r_2+r_2z}{1-r_2}\right)^{\frac{\gamma_j}{T}n}}{1-\left(\frac{1-r_2+r_2z}{1-r_2}\right)^n}. \quad (3.8)$$

If $\frac{\gamma_j}{T} < i \leq \frac{\delta_j}{T}$:

$$P^{\mathrm{uni}}\big((i-1)n+1 \leq u-v \leq in\big) = \sum_{l=(i-1)n+1}^{in} q_j \left[1-(1-r_2)^{\frac{\delta_j}{T}n-l} \right] =$$

$$= q_j n - q_j (1-r_2)^{\frac{\delta_j}{T}n} \frac{1-(1-r_2)^n}{r_2(1-r_2)^n} \frac{1}{(1-r_2)^{(i-1)n}},$$

and the generating function of entering customers is

$$\sum_{i=\frac{\gamma_j}{T}+1}^{\frac{\delta_j}{T}} \left[q_j n - q_j (1-r_2)^{\frac{\delta_j}{T}n} \frac{1-(1-r_2)^n}{r_2(1-r_2)^n} \frac{1}{(1-r_2)^{(i-1)n}} \right] (1-r_2+r_2z)^{in} =$$

$$= q_j n (1-r_2+r_2z)^n \frac{(1-r_2+r_2z)^{\frac{\gamma_j}{T}n} - (1-r_2+r_2z)^{\frac{\delta_j}{T}n}}{1-(1-r_2+r_2z)^n} -$$

$$-q_j (1-r_2)^{\frac{\delta_j}{T}n} \frac{1-(1-r_2)^n}{r_2(1-r_2)^n} (1-r_2+r_2z)^n \frac{\left(\frac{1-r_2+r_2z}{1-r_2}\right)^{\frac{\gamma_j}{T}n} - \left(\frac{1-r_2+r_2z}{1-r_2}\right)^{\frac{\delta_j}{T}n}}{1-\left(\frac{1-r_2+r_2z}{1-r_2}\right)^n}. \quad (3.9)$$

The probability that the waiting time is zero (which happens when the next customer enters during the last time-slice of the service of the previous one) is

$$\sum_{k=\frac{\gamma_j}{T}n+1}^{\frac{\delta_j}{T}n} q_j (1-r_2)^{k-1} r_2 = q_j \left[(1-r_2)^{\frac{\gamma_j}{T}n} - (1-r_2)^{\frac{\delta_j}{T}n} \right], \quad (3.10)$$

while the probability that there is no entry at all is

$$\sum_{k=\frac{\gamma_j}{T}n+1}^{\frac{\delta_j}{T}n} q_j\,(1-r_2)^k = \frac{q_j}{r_2}\left[(1-r_2)^{\frac{\gamma_j}{T}n+1} - (1-r_2)^{\frac{\delta_j}{T}n+1}\right]. \tag{3.11}$$

Bearing in mind that we examined those possibilities when a customer obligatorily enters, generating functions (3.8), (3.9), and (3.10) have to be multiplied by z, then added to (3.11), which yields (3.2).

If the third collision treatment method is applied, then $A_{12}(z)$ has to be determined. In this case the service starts with the simultaneous arrival of two customers of different types, and the frst-type one is served first. Thus, the duration of its service is l units with probability $q_1^{l-1}\,(1-q_1)$. The probability that it is in $\overline{(i-1)\,n+1, in}$ and hence, the probability that the waiting time of the next customer is in in the case of geometrically distributed service time is

$$P^{\text{geo}}\big(\,(i-1)\,n+1 \le u \le in\big) = \sum_{l=(i-1)n+1}^{in} q_1^{l-1}\,(1-q_1) = q_1^{(i-1)n} - q_1^{in}.$$

In the case of uniform distribution the service time can take on any value between $\frac{\gamma_1}{T}n+1$ and $\frac{\delta_1}{T}n$ with equal probability $q_1 = \frac{T}{n(\delta_1-\gamma_1)}$, so

$$P^{\text{uni}}\big(\,(i-1)\,n+1 \le u \le in\big) = \sum_{l=(i-1)n+1}^{in} q_1 = q_1 n$$

for all $\frac{\gamma_1}{T} < i \le \frac{\delta_1}{T}$.

Taking into account that one customer of second type is already present at the start of the service of the first-type customer, the generating functions are

$$A_{12}^{\text{geo}}(z) = z\sum_{i=1}^{\infty}\left[q_1^{(i-1)n} - q_1^{in}\right](1-r_2+r_2z)^{in} = z\frac{(1-q_1^n)\,(1-r_2+r_2z)^n}{1-q_1^n\,(1-r_2+r_2z)^n},$$

and

$$A_{12}^{\mathrm{uni}}(z) = z \sum_{i=\frac{\gamma_1}{T}+1}^{\frac{\delta_1}{T}} nq_1 \left(1 - r_2 + r_2 z\right)^{in} =$$

$$= znq_1 \frac{\left(1 - r_2 + r_2 z\right)^{\left(\frac{\gamma_1}{T}+1\right)n} - \left(1 - r_2 + r_2 z\right)^{\left(\frac{\delta_1}{T}+1\right)n}}{1 - \left(1 - r_2 + r_2 z\right)^n},$$

which are identical with (3.3) and (3.4).

The probability of appearance of at least one customer of any type in a time-slice is

$$1 - (1 - r_1)(1 - r_2) = r_1 + r_2 - r_1 r_2.$$

The busy period can start with the arrival of a first-type customer alone with probability $\frac{r_1(1-r_2)}{r_1+r_2-r_1r_2}$, with the arrival of a single second-type customer with probability $\frac{r_2(1-r_1)}{r_1+r_2-r_1r_2}$, and with both of them with probability $\frac{r_1r_2}{r_1+r_2-r_1r_2}$, these easily explain (3.7c) (collision discipline III.). In the case of collision treatment method I., customers of both types are lost if they arrive during the same time-slice, and this may be interpreted as a service of zero length (the system stays in free state with this probability), which results in the generating function (3.7a). If collision discipline II. is applied, then the busy period can start with the service of a first-type customer with probability $\frac{r_1}{r_1+r_2-r_1r_2}$ (no matter whether there was a refused second-type one at the same time), and starts with the service of a second-type customer with probability $\frac{r_2(1-r_1)}{r_1+r_2-r_1r_2}$, which gives reason to (3.7b).

Finally, we are going to determine transition probabilities b_i. In this case when the service of the actual customer begins, the next one is already present. Let $x = u - \left\lfloor \frac{u-1}{n} \right\rfloor n$, i. e. x is the service time mod n ($1 \le x \le n$), and let y denote the interarrival time mod n ($1 \le y \le n$). The time elapsed between the starting moments of services of these two consecutive customers will be

$$t_0 = \begin{cases} \left\lfloor \frac{u-1}{n} \right\rfloor n + y, & \text{if } x \le y; \\ \left(\left\lfloor \frac{u-1}{n} \right\rfloor + 1\right) n + y, & \text{if } x > y. \end{cases}$$

Now take y fixed, and consider the usual set of integers $\overline{in+1, (i+1)\,n}$. If the service is completed until $in + y$ (inclusive), then the time in question is $in + y$. The probability of this event in the case of geometric distribution is

$$\sum_{k=in+1}^{in+y} q_2^{k-1}\left(1-q_2\right) = q_2^{in} - q_2^{in+y};$$

and it is yq_2 in the case of uniformly distributed service time. If the service finishes later than $in + y$, then the difference of instants between starting services is $(i+1)\,n + y$ with probability

$$\sum_{k=in+y+1}^{(i+1)n} q_2^{k-1}\left(1-q_2\right) = q_2^{in+y} - q_2^{(i+1)n}$$

in the case of geometric; and $(n-y)\,q_2$ in the case of uniform distribution. Summation has to be extended over all possible values of service times, therefore, the generating function of the number of entering customers on condition that interarrival time $\bmod n$ is equal to y in the case of geometric service time distribution is

$$\begin{aligned}\sum_{i=0}^{\infty} \Big[&\left(q_2^{in} - q_2^{in+y}\right)\left(1-r_2+r_2z\right)^{in+y} + \\ &+ \left(q_2^{in+y} - q_2^{(i+1)n}\right)\left(1-r_2+r_2z\right)^{(i+1)n+y} \Big] = \\ &= \frac{\left(1-q_2^y\right)\left(1-r_2+r_2z\right)^y}{1-q_2^n\left(1-r_2+r_2z\right)^n} + \frac{\left(q_2^y-q_2^n\right)\left(1-r_2+r_2z\right)^{n+y}}{1-q_2^n\left(1-r_2+r_2z\right)^n};\end{aligned}$$

and in the case of uniformly distributed service time

$$\begin{aligned}\sum_{i=\frac{\gamma_2}{T}}^{\frac{\delta_2}{T}-1} &\left[q_2 y\left(1-r_2+r_2z\right)^{in+y} + q_2\left(n-y\right)\left(1-r_2+r_2z\right)^{(i+1)n+y} \right] = \\ &= q_2\left(1-r_2+r_2z\right)^y \left[y + (n-y)\left(1-r_2+r_2z\right)^n \right] \times \\ &\times \frac{\left(1-r_2+r_2z\right)^{\frac{\gamma_2}{T}n} - \left(1-r_2+r_2z\right)^{\frac{\delta_2}{T}n}}{1-\left(1-r_2+r_2z\right)^n}.\end{aligned}$$

The random variable of y has truncated geometric distribution with probabilities $\frac{(1-r_2)^{k-1}r_2}{1-(1-r_2)^n}$ $(k = 1, 2, \ldots, n)$. Previously calculated sums have to be multiplied by $\frac{(1-r_2)^{y-1}r_2}{1-(1-r_2)^n}$, and summed up with respect to y from 1 to n. Expanding these sums we finally receive (3.5) and (3.6). ∎

As far as the fomulae are concerned, the equilibrium distribution of this process is the same as in the continuous case. However, the final result depends on the generating functions and their derivatives, so we partly repeat the statement of Theorem 2.1.2.

Theorem 3.1.2 *The generating function of the equilibrium distribution of this chain is:*

$$P(z) = \sum_{i=0}^{\infty} p_i z^i = \frac{p_0\big(zC(z) - B(z)\big) + p_1 z\big(A_2(z) - B(z)\big)}{z - B(z)}, \tag{3.12}$$

where p_0 and p_1 are the first two probabilities of the equilibrium distribution. They are connected with the relation $p_1 = \frac{1-c_0}{a_{20}} p_0$, and

$$p_0 = \frac{1 - B'(1)}{1 - B'(1) + C'(1) + \frac{1-c_0}{a_{20}}\big(A_2'(1) - B'(1)\big)}, \tag{3.13}$$

where

$$A_j'^{\,geo}(1) = \frac{r_2}{1 - q_j(1 - r_2)}\left(1 + \frac{n r_2 q_j}{1 - q_j^n}\right),$$

$$A_j'^{\,uni}(1) = -a_{j0} + \frac{T}{\delta_j - \gamma_j}\left[(1 - r_2)^{\frac{\gamma_j}{T}n} - (1 - r_2)^{\frac{\delta_j}{T}n}\right]\frac{(1 - r_2)^n}{1 - (1 - r_2)^n} + \\ + \frac{n r_2}{T}\frac{\gamma_j + \delta_j + T}{2},$$

$$A_{12}'^{\,geo}(1) = 1 + \frac{n r_2}{1 - q_1^n},$$

$$A_{12}'^{\,uni}(1) = 1 + \frac{n r_2}{T}\frac{\gamma_1 + \delta_1 + T}{2},$$

$$B'^{\,geo}(1) = 1 - \frac{n r_2}{1 - (1 - r_2)^n}\left[(1 - r_2)^n - \frac{r_2 q_2}{1 - q_2^n}\frac{1 - q_2^n(1 - r_2)^n}{1 - q_2(1 - r_2)}\right], \tag{3.14a}$$

$$B'^{\,uni}(1) = \frac{n r_2}{T}\frac{\gamma_2 + \delta_2 + T}{2}, \tag{3.14b}$$

and $C'(1)$ depends on collision policies:

I. $C'(1) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2} A_1'(1) + \frac{r_2(1-r_1)}{r_1+r_2-r_1r_2} A_2'(1)\,,$

II. $C'(1) = \frac{r_1}{r_1+r_2-r_1r_2} A_1'(1) + \frac{r_2(1-r_1)}{r_1+r_2-r_1r_2} A_2'(1)\,,$

III. $C'(1) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2} A_1'(1) + \frac{r_2(1-r_1)}{r_1+r_2-r_1r_2} A_2'(1) + \frac{r_1r_2}{r_1+r_2-r_1r_2} A_{12}'(1)\,.$

Lemma 3.1.3 *Expression $C'(1) + \frac{1-c_0}{a_{20}}\big(A_2'(1) - B'(1)\big)$ is always positive for any values of the parameters, for any combination of service time distributions and for any collision discipline.*

Proof If the service time of second-type customers is geometrically distributed, then after several factorizations we get

$$A_2'(1) - B'(1) = -\frac{(1-r_2)(1-q_2)}{1-q_2(1-r_2)} + \frac{nr_2(1-r_2)^n}{1-(1-r_2)^n} \cdot \frac{1-q_2}{1-q_2(1-r_2)}.$$

Considering that $a_{20}^{\text{geo}} = \frac{(1-r_2)(1-q_2)}{1-q_2(1-r_2)}$,

$$\frac{1-c_0}{a_{20}}\big(A_2'(1) - B'(1)\big) = (1-c_0)\left(-1 + \frac{nr_2}{1-r_2} \cdot \frac{(1-r_2)^n}{1-(1-r_2)^n}\right).$$

If the service time of second type customers is uniformly distributed,

$$\frac{1-c_0}{a_{20}}\big(A_2'(1) - B'(1)\big) =$$
$$= (1-c_0)\left[-1 + \frac{T}{a_{20}(\delta_2-\gamma_2)}\left((1-r_2)^{\frac{\gamma_2}{T}n} - (1-r_2)^{\frac{\delta_2}{T}n}\right)\frac{(1-r_2)^n}{1-(1-r_2)^n}\right].$$

Substituting $a_{20}^{\text{uni}} = \frac{T}{n(\delta_2-\gamma_2)}\frac{1-r_2}{r_2}\left((1-r_2)^{\frac{\gamma_2}{T}n} - (1-r_2)^{\frac{\delta_2}{T}n}\right)$ in the formula, we get:

$$\frac{1-c_0}{a_{20}}\big(A_2'(1) - B'(1)\big) = (1-c_0)\left(-1 + \frac{nr_2}{1-r_2} \cdot \frac{(1-r_2)^n}{1-(1-r_2)^n}\right);$$

which is the same as in the previous case.

Finally, $C'(1)$ will be transformed in the following way:

$$C'(1) = \sum_{i=0}^{\infty} ic_i = \sum_{i=0}^{\infty} c_i - c_0 + \sum_{i=2}^{\infty}(i-1)\,c_i = 1 - c_0 + \sum_{i=2}^{\infty}(i-1)\,c_i.$$

Substituing all in, $1 - c_0$ cancels in both cases:

$$C'(1) + \frac{1-c_0}{a_{20}}\big(A_2'(1) - B'(1)\big) =$$
$$= \sum_{i=2}^{\infty} (i-1)\, c_i + (1-c_0)\, \frac{nr_2}{1-r_2} \frac{(1-r_2)^n}{1-(1-r_2)^n}.$$

Rewritten in this form it is obvious that $C'(1) + \frac{1-c_0}{a_{20}}\big(A_2'(1) - B'(1)\big) > 0$.■

Theorem 3.1.4 *The condition of existence of the equilibrium distribution is the fulfilment of one of the following inequalities.*

If the service time of second-type customers is geometrically distributed (regardless of the distribution of service time of first-type customers and the applied collision discipline):

$$\frac{r_2 q_2}{1-q_2^n} \frac{1 - q_2^n (1-r_2)^n}{1 - q_2 (1-r_2)} < (1-r_2)^n . \tag{3.15}$$

If the service time of second-type customers is uniformly distributed (regardless of the distribution of service time of first-type customers and collision discipline):

$$\frac{nr_2}{T} \frac{\gamma_2 + \delta_2 + T}{2} < 1. \tag{3.16}$$

Proof As the embedded Markov-chain is irreducible and aperiodic, the process is ergodic iff $0 < p_0 < 1$. Applying Theorem 3.1.2 and Lemma 3.1.3, $p_0 = \frac{1-B'(1)}{1-B'(1)+K}$, where K is a positive constant. Thus, the condition simplifies into

$$1 - B'(1) > 0,$$

which – together with (3.14a) and (3.14b) – gives (3.15) and (3.16). ■

One remarkable thing about (3.15) and (3.16) is that they depend neither on first-type customers nor on collision policies, i. e. these have no effect on the ergodicity of the process. Moreover, the formula expressing the condition in the uniform case can be interpreted easily: considering that $\frac{T}{n}\frac{1}{r_2}$ is the average inter-arrival time, the condition rewritten in the form

$$\frac{\gamma_2 + \delta_2}{2} + \frac{T}{2} < \frac{T}{nr_2}$$

necessitates that the average service time increased by the average idle time (on average $\frac{T}{2}$ time is needed for the next customer in the queue to reach the starting position) be less than the average inter-arrival time.

The dependence on input parameters of condition (3.16) is simple, but it is more complex in (3.15). For the latter case it is more convenient if parameters that guarantee ergodicity are represented in a graph. Numerical examination of inequality (3.15) reveals that for fixed values of n and q_2, the process is ergodic if $r_2 \in (0, r_2^{\max})$. The shaded areas (the areas under the curves, excluding the curves themselves) in Figure 3.1 show pairs (q_2, r_2) suitable for ergodicity at some arbitrary values of n (at $n = 1$, 2, 4, 6, 8, and 10).

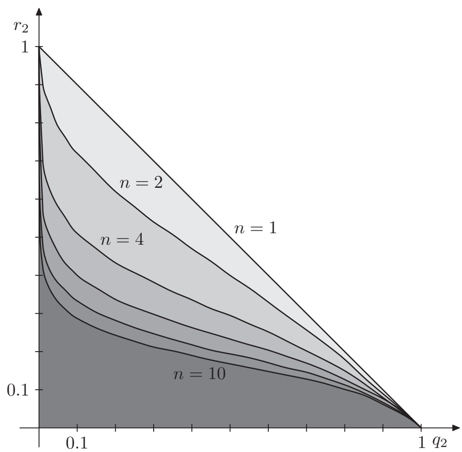


Figure 3.1: *Pairs (q_2, r_2) suitable for ergodicity*

It can be seen that in the (somewhat degenerate) case $n = 1$, the condition of ergodicity is $r_2 < 1 - q_2$, which has a clear probabilistic interpretation; namely, that the probability of appearance of a new customer of second-type during a time slice (which is T in this case) has to be less than the probability of completing a service during the same amount of time.

3.2 Validating results

Equilibrium probabilities of states 0 and 1

In Theorem 3.1.2 the equilibrium probability of free state was determined as the function of the derivatives of the generating functions while $z \to 1$, which were given, as well. Probability p_1 was also given by a simple relation to p_0, whereas other equilibrium probabilities can be derived from (3.12).

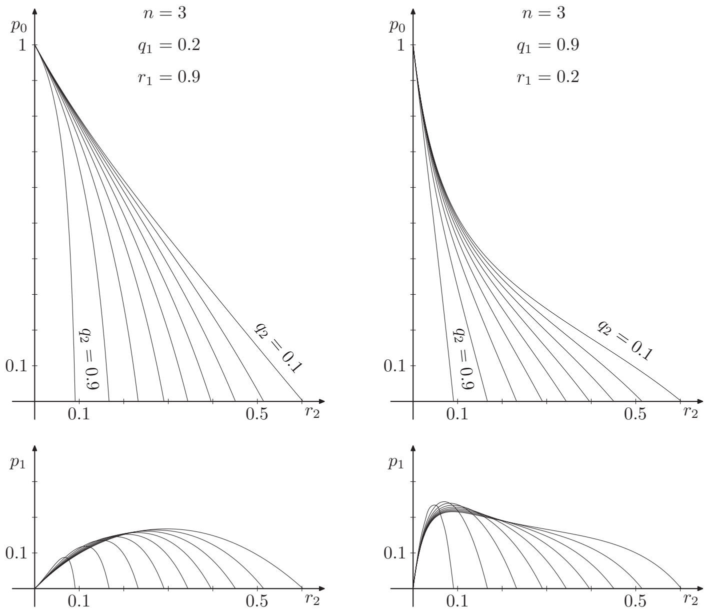


Figure 3.2: *Probabilities p_0 and p_1 with geometrically distributed service times and collision discipline I.*

Similarly to the continuous case, it is advisable to have a closer look at the first two, explicitly given probabilities, with special attention to how they vary with the change of the input parameters. Considering the high num-

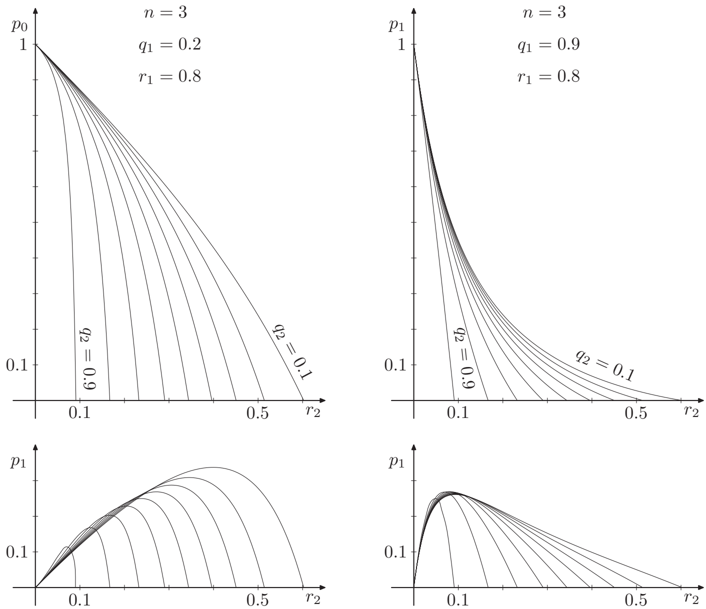


Figure 3.3: *Probabilities p_0 and p_1 with geometrically distributed service times and collision discipline II.*

ber of input parameters and our extra assumption on the boundaries of the uniform distributions, this seems feasible only with the service time distribution of second-type customers being exponential. Figures 3.2, 3.3 and 3.4 show their behaviour when service times are exponentially distributed, and all three collision treatment policies are considered, respectively. The load intensity of second-type customers is chosen as the main independent variable at nine different values of their service capacity ($q_2 = 0.1, 0.2, \ldots, 0.9$). Two sets of parameters of first-type customers are chosen to have the opportunity to compare the effect of a weak load on a high-capacity system, and a heavy load on a low-capacity one.

As it is expected, observations similar to the continuous case can be made.

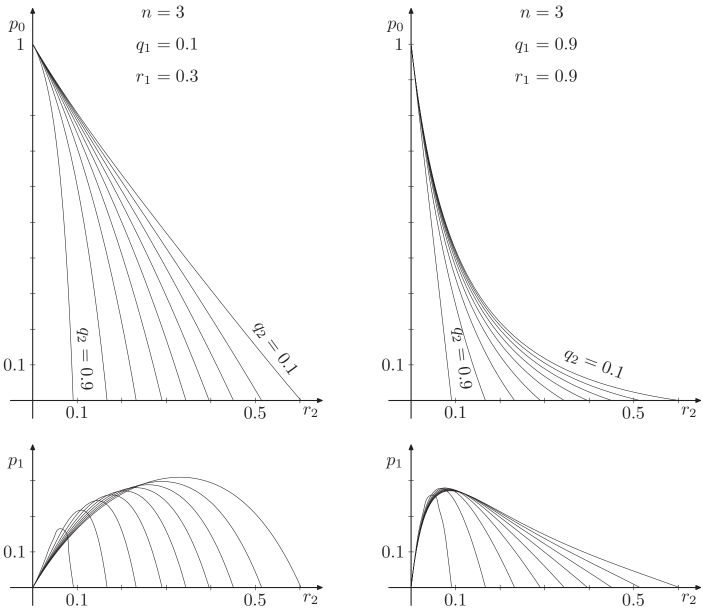


Figure 3.4: *Probabilities p_0 and p_1 with geometrically distributed service times and collision discipline III.*

Obviously, the probability of free state is a decreasing function of the input intensity; whereas p_1 has a maximum between 0 and $r_2^{\max}$, where $r_2^{\max}$ is the largest possible value of the input intensity allowed by ergodicity condition (3.15). With chosen parameters ($n = 3$ and $q_2 = 0.1, 0.2, \ldots, 0.9$) these are 0.60, 0.51, 0.45, 0.39, 0.34, 0.29, 0.23, 0.16, 0.09 (truncated to the second significant decimal), respectively. Values of $r_2^{\max}$ for other parameters have been seen previously in Figure 3.1, and can be found later in Table 3.1.

Naturally, although first-type customers do have a considerable effect on the first two equilibrium probabilities (they decrease at lower values of r_2 if first-type load is heavier), but they have no effect on the ergodicity of the process; all functions take their zero-points at the same values. This is in accordance with (3.15), as it contains no first-type parameters.

Queue-length

Dealing with queuing systems an exciting question always emerges, namely what the mean value of the queue-length is, or how it depends on the parameters, especially on the intensity of the input with other parameters fixed at certain values.

In this section we give the formula for the expected value of the queue-length, which holds for any type of queuing systems with two types of customers examined so far. In the discrete case when service time distributions are geometric, the queue-length is given explicitly, and its dependence on input parameters is analyzed graphically.

Theorem 3.2.1 *The expected value of the length of the queues in cyclic-waiting systems serving two types of customers is given by*

$$P'(1) = 1 - p_0 + \frac{(1-p_0)\,B''(1) + p_0\left(C''(1) + \frac{1-c_0}{a_{20}}\left(A_2''(1) - B''(1)\right)\right)}{2\left(1 - B'(1)\right)},$$

where p_0 *is given by* (3.13), $B'(1)$ *is given by* (3.14a),

$$A_j''^{\,geo}(1) = \frac{nr_2^2 q_j\left(2 + nr_2 - r_2\right)}{\left(1 - q_j^n\right)\left(1 - q_j\left(1 - r_2\right)\right)} + \frac{2n^2 r_2^3 q_j^{n+1}}{\left(1 - q_j^n\right)^2\left(1 - q_j\left(1 - r_2\right)\right)},$$

$$A_{12}''^{\,geo}(1) = \frac{nr_2\left(2 + nr_2 - r_2\right)}{1 - q_1^n} + \frac{2n^2 r_2^2 q_1^n}{\left(1 - q_1^n\right)^2},$$

$$\begin{aligned} B''^{\,geo}(1) = {} & 2\left(1 - r_2\right) - \frac{nr_2\left(1 - r_2\right)^n\left(2 + nr_2 - r_2\right)}{1 - \left(1 - r_2\right)^n} + \\ & + \frac{nr_2^3 q_2}{\left(1 - \left(1 - r_2\right)^n\right)\left(1 - q_2^n\right)} \frac{1 - q_2^n\left(1 - r_2\right)^n}{1 - q_2\left(1 - r_2\right)}\left(n - 1 + \frac{2}{1 - q_2\left(1 - r_2\right)}\right) + \\ & + \frac{2n^2 r_2^3 q_2^{n+1}}{\left(1 - q_2^n\right)^2\left(1 - q_2\left(1 - r_2\right)\right)}, \end{aligned}$$

if $n \geq 2$; and $C''(1)$ depends on the collision treatment policy:

I. $C''(1) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2} A_1''(1) + \frac{r_2(1-r_1)}{r_1+r_2-r_1r_2} A_2''(1)$,

II. $C''(1) = \frac{r_1}{r_1+r_2-r_1r_2} A_1''(1) + \frac{r_2(1-r_1)}{r_1+r_2-r_1r_2} A_2''(1)$,

III. $C''(1) = \frac{r_1(1-r_2)}{r_1+r_2-r_1r_2} A_1''(1) + \frac{r_2(1-r_1)}{r_1+r_2-r_1r_2} A_2''(1) + \frac{r_1r_2}{r_1+r_2-r_1r_2} A_{12}''(1)$.

Before depicting the function of the queue-length it is essential to know for what parameters the system has an equilibrium distribution. Table 3.1 shows values of $r_2^{\max}$ (truncated to the second decimal digit) in the case of some given values of n and q_2.

Table 3.1: *Maximum values of* r_2 *(approx.) for ergodicity*

		q_2								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
n	2	0.72	0.62	0.54	0.47	0.40	0.33	0.26	0.18	0.09
	3	0.60	0.51	0.45	0.39	0.34	0.29	0.23	0.16	0.09
	4	0.51	0.43	0.38	0.34	0.29	0.25	0.20	0.15	0.08
	5	0.45	0.38	0.33	0.29	0.26	0.22	0.19	0.14	0.08
	6	0.40	0.34	0.30	0.26	0.23	0.20	0.17	0.13	0.08
	7	0.36	0.31	0.27	0.24	0.21	0.19	0.16	0.12	0.07
	8	0.33	0.28	0.25	0.22	0.20	0.17	0.15	0.11	0.07
	9	0.31	0.26	0.23	0.20	0.18	0.16	0.14	0.11	0.07
	10	0.29	0.24	0.21	0.19	0.17	0.15	0.13	0.10	0.06

On the following pages graphs are presented showing the behaviour of the mean length of the specified queue as a function of r_2 at nine given values of q_2 (from 0.1 upto 0.9) and at arbitrary values of n, r_1 and q_1. Just exactly how it is expected, parameters r_1 and q_1 have very weak effect on the queue-length. Although queues are longer on average, if first-type load is heavier (the curves of queue-lengths are steeper with increasing second-type load), but they have no effect at all on the ergodicity of the process; the curves have their vertical asymptotes at the same values of r_2 for fixed values of q_2.

Collision policy I: In the case of collision both customers are refused.

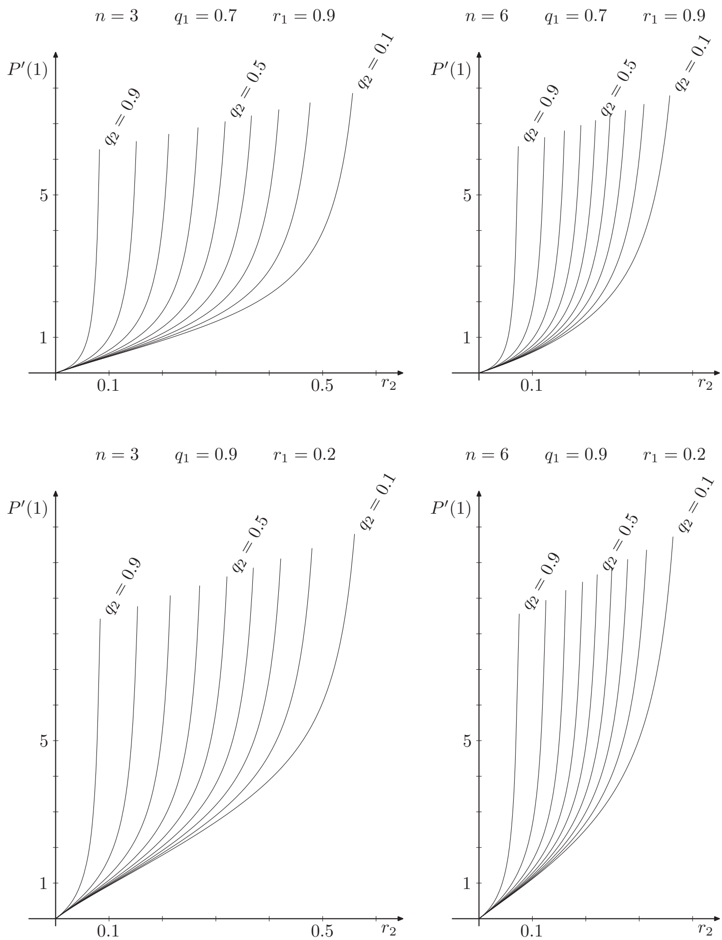


Figure 3.5: *Queue-length with collision policy I.*

Collision policy II: In the case of collision first-type customers are accepted but second-type ones are refused.

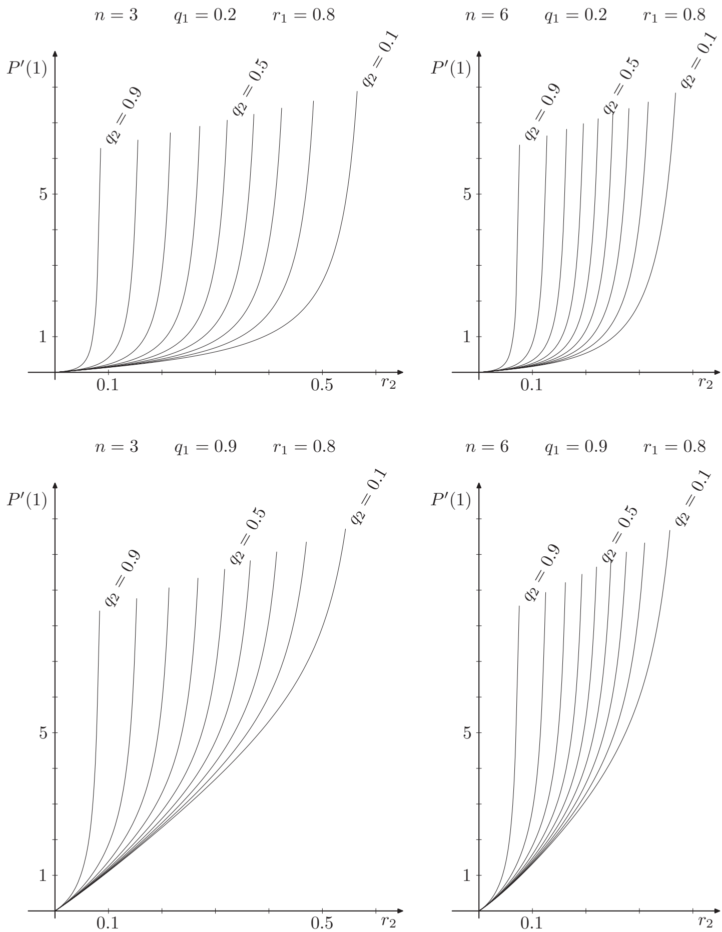


Figure 3.6: *Queue-length with collision policy II.*

Collision policy III: In the case of collision customers of both types are accepted, and first-type ones are served first.

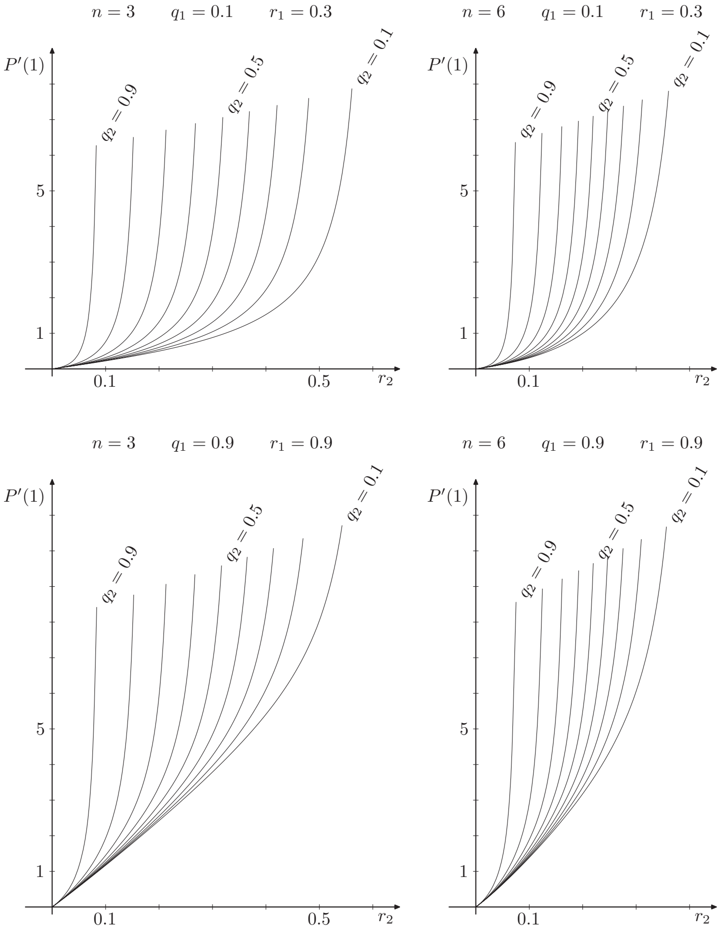


Figure 3.7: *Queue-length with collision policy III.*

Chapter 4

Queuing systems with vacation

To determine the equilibrium distribution of different queuing systems, the embedded Markov-chain technique is often used, which leads to the Pollaczek–Khinchin formula. This gives the solution of the problem, probabilities can be derived from it by means of differentiation, but it gives very complicated results. These difficulties induced the search for other methods. In [4] Brière and Chaudhry considered bulk-arrival systems, and found a recursive algorithm for different service time distributions. Lakatos used another approach, he described queuing systems with the help of Kovalenko's piece-wise linear processes. This makes it possible to calculate the desired probabilities on the basis of the mean length of a busy period and the mean value of time spent in different states. In [57] he gave recursive formulae for the equilibrium distribution of the ordinary $M/G/1$ system, and for the system where after each busy period there is a vacation.

4.1 Bulk-arrival systems with vacation at the beginning of the busy period

In [64] Lakatos investigated such a queuing system where the arrival of the first customer initiates a vacation, and service can only start when the system is prepared for it during the vacation. This type of system fits the conditions of regenerative processes, and so it is quite natural to use the results and tools of these processes to describe them. This and the above-mentioned method made it possible to give recursive formulae for the equilibrium distribution of this type of system.

In this chapter we are going to generalize results for the latter system, which accepts bulk-arrivals.

Notations

We introduce the following notations:

λ – arrival rate of groups of customers;

$B(x)$ – the distribution function of the service time of a customer, $b(s) = \int_0^\infty e^{-sx}\,\mathrm{d}B(x)$ its Laplace-Stieltjes transform and $\tau = \int_0^\infty x\,\mathrm{d}B(x)$ its mean value;

$c_k = \int_0^\infty \frac{(\lambda x)^k}{k!} e^{-\lambda x}\,\mathrm{d}B(x)$ – the probability of appearance of k new groups of customers during the service of a customer;

$\varrho = \alpha\lambda\tau$ – the utilization factor of the system;

$P_i(k)$ – the probability of event that in i groups k requests appear altogether;

$\Gamma_1(s)$ – the Laplace-Stieltjes transform of the busy period's distribution function in the ordinary (non-vacational) system if it starts with the entry of one customer;

$\Gamma(s)$ – the Laplace-Stieltjes transform of the busy period's distribution function;

$\zeta = \frac{\alpha\tau}{1-\varrho}$ – the mean value of duration of the busy period in the ordinary (non-vacational) bulk-arrival $M/G/1$ system;

ζ' – the mean value of duration of the busy period in the system with vacation;

ζ_i – the mean value of time spent above the i^{th} level for a busy period in the system without vacation;

ζ_i' – the mean value of time spent above the i^{th} level for a busy period in the system with vacation;

ξ_i – the mean value of time spent on the i^{th} level for a busy period in the system without vacation;

ξ_i' – the mean value of time spent on the i^{th} level for a busy period in the system with vacation;

$D(x)$ – the distribution function of vacation, $d(s) = \int_0^\infty e^{-sx}\,\mathrm{d}D(x)$ its Laplace-Stieltjes transform and $\eta = \int_0^\infty x\,\mathrm{d}D(x)$ its mean value;

$d_k = \int_0^\infty \frac{(\lambda x)^k}{k!} e^{-\lambda x}\,\mathrm{d}D(x)$ – the probability of k new entries during vacation;

$G(z) = \sum_{k=1}^\infty g_k z^k$ – the generating function of number of requests in the arriving group, $\alpha = G'(1)$ its mean value;

$E(z) = \sum_{k=0}^\infty e_k z^k$ – the generating function of number of requests appearing during vacation;

$F(z) = \sum_{k=1}^\infty f_k z^k$ – the generating function of number of requests being present at the end of vacation.

4.2 Results

Theorem 4.2.1 *In the bulk-arrival $M/G/1$ system with vacation at the beginning of the busy period the ergodic distribution exists if $\varrho < 1$; and equilibrium probabilities are determined by*

$$p_i = \frac{\xi_i'}{\zeta'}, \qquad i = 0, 1, 2, \ldots$$

PROOF The theorem is a direct consequence of [77, Theorem 1.7.1], which states that the long-run fraction of arrivals who find the system in the set B of states is equal to the long-run fraction of time the system is in the set B of states. This statement is usually abbreviated as PASTA (Poisson arrivals see time averages). ∎

Lemma 4.2.2 *The mean number of entering customers during vacation; and the mean number of customers present at the end of vacation are*

$$E'(1) = \alpha\lambda\eta, \qquad \textit{and} \qquad F'(1) = \alpha + \alpha\lambda\eta,$$

respectively.

Proof For the generating function of customers entering during vacation we get

$$
\begin{aligned}
E(z) = \int_0^\infty e^{-\lambda x}\,\mathrm{d}D(x) + \int_0^\infty \sum_{i=1}^\infty \sum_{k=i}^\infty e^{-\lambda x}\frac{(\lambda x)^i}{i!}P_i(k)z^k\,\mathrm{d}D(x) = \\
= \int_0^\infty e^{-\lambda x}\sum_{i=0}^\infty \frac{(\lambda x)^i}{i!}\sum_{k=i}^\infty P_i(k)z^k\,\mathrm{d}D(x) = \\
= \int_0^\infty e^{-\lambda x}\sum_{i=0}^\infty \frac{(\lambda x G(z))^i}{i!}\,\mathrm{d}D(x) = d\left(\lambda\left(1-G(z)\right)\right).
\end{aligned}
$$

From this $E'(1) = \left[d\left(\lambda\left(1-G(z)\right)\right)\right]'\Big|_{z=1} = \alpha\lambda\eta$.

By definition of generating functions $F(z)$, $G(z)$ and $E(z)$ it is obvious that $F(z) = G(z)E(z)$, thus $F'(1) = G'(1) + E'(1) = \alpha + \alpha\lambda\eta$. ■

Lemma 4.2.3 *The mean value of duration of the busy period is*

$$
\zeta' = \frac{\eta + \alpha\tau}{1-\varrho}. \tag{4.1}
$$

Proof In [58] it is shown that if the busy period starts with the appearance of one customer then the Laplace-Stieltjes transform of its distribution function $\Gamma_1(s)$ satisfies the functional equation

$$
\Gamma_1(s) = b\left(s + \lambda - \lambda G(\Gamma_1(s))\right).
$$

From this we obtain the mean value of the duration of the busy period (starting with one request) in the ordinary system: $-\Gamma_1'(0) = \frac{\tau}{1-\alpha\lambda\tau}$.

We determine the Laplace-Stieltjes transform of the distribution function of the busy period. The busy period begins with the entry of the first group, which initiates a vacation. The busy period consists of this vacation and the service of all customers entering during it and during the service of the subsequent ones. The corresponding transforms are $d\left(s + \lambda - \lambda G(\Gamma_1(s))\right)$ for the service of customers entering during vacation and $G(b\left(s + \lambda - \lambda G(\Gamma_1(s))\right))$

for the customers generated by the first group. So, the desired Laplace-Stieltjes transform is

$$\Gamma(s) = d\left(s + \lambda - \lambda G(\Gamma_1(s))\right) G(b\left(s + \lambda - \lambda G(\Gamma_1(s))\right)).$$

We get the mean value of duration of the busy period through differentiation, and using the previously obtained formula for $\Gamma_1'(0)$, it equals

$$\zeta' = -\Gamma(0) = \frac{\eta + \alpha\tau}{1 - \varrho}.$$

■

Theorem 4.2.4 *The mean value of time spent on the i^{th} level for a busy period satisfies the recurrence relations*

$$\begin{aligned}
\xi_0' &= \tau, \\
\xi_1' &= \frac{\tau}{c_0} + f_1\left(\eta - \tau\right), \\
\xi_2' &= \xi_2 + (2 - f_1)\left(\frac{\tau}{1-\varrho} - \zeta_1\right) - \frac{\tau}{c_0} + f_2\left(\eta - \tau\right), \\
&\vdots \\
\xi_i' &= \xi_i + \sum_{k=2}^{i-1}\left(1 - f_1 - \ldots - f_{i-k}\right)\xi_k + \\
&+ \left(1 - f_1 - \ldots - f_{i-1}\right)\left(\frac{\tau}{1-\varrho} - \zeta_1\right) + f_i\left(\eta - \tau\right), \qquad (i \geq 3).
\end{aligned}$$

Proof By the definition of the embedded Markov-chain, the only case when our system is on level 0 is when the very last request is being serviced, the mean value of time spent in state 0 is obviously τ.

We determine the mean value of the period when there is one request in the system. This will give $\xi_0' + \xi_1'$, as the service of the last customer belongs to state 0. We are going to consider three different cases.

If the process begins with the entry of one customer and no other customers arrive during vacation, then according to [58] the desired mean value is

$$g_1 d_0 \left(\eta + \frac{\tau}{c_0}\right).$$

If the process begins with the entry of one customer and there is at least one entry during the vacation the desired mean value is

$$g_1 (1 - d_0) \left(\tau + \frac{\tau}{c_0} \right).$$

If the process begins with the entry of a group of more than one request, then the desired mean value is

$$(1 - g_1) \left(\tau + \frac{\tau}{c_0} \right).$$

The sum of these three terms give

$$\xi_0' + \xi_1' = \tau + \frac{\tau}{c_0} + g_1 d_0 (\eta - \tau) = \tau + \frac{\tau}{c_0} + f_1 (\eta - \tau).$$

We determine the mean value of time spent above the first level for a busy period. We distinguish between two cases. If at the end of the vacation there is only one request present, then, by [58], this value is

$$f_1 \tau \left(\frac{1}{1 - \varrho} - \frac{1}{c_0} \right).$$

If at the end of the vacation there are more than one requests present, then the mean value in question is

$$\sum_{k=2}^{\infty} f_k \left(\eta + (k-1) \frac{\tau}{1 - \varrho} - \tau + \tau \left(\frac{1}{1 - \varrho} - \frac{1}{c_0} \right) \right) =$$
$$= (1 - f_1) \left(\eta - \tau - \frac{\tau}{c_0} \right) + \frac{\tau}{1 - \varrho} (\alpha + \alpha \lambda \eta - f_1).$$

The sum of these two terms gives ζ_1':

$$\zeta_1' = (1 - f_1)(\eta - \tau) - \frac{\tau}{c_0} + \frac{\alpha \tau}{1 - \varrho} (1 + \lambda \eta).$$

In order to check the correctness of obtained values, we add them (using $f_1 = g_1 e_0 = g_1 d_0$):

$$\xi_0' + \xi_1' + \zeta_1' = \tau + \frac{\tau}{c_0} + g_1 d_0 (\eta - \tau) +$$
$$+ (1 - f_1)(\eta - \tau) - \frac{\tau}{c_0} + \frac{\alpha \tau}{1 - \varrho} (1 + \lambda \eta) = \frac{\eta + \alpha \tau}{1 - \varrho},$$

which coincides with (4.1).

Next we find the mean value of time spent above second level. We have several possibilities depending on how many customers are present at the end of the vacation. If there is only one, then we have

$$f_1\zeta_2.$$

If at the end of the vacation there are two customers in the system, spending ζ_1 above second level we come to the first one, then spending ζ_2 above second level the busy period ends, the corresponding value is

$$f_2\left(\zeta_1+\zeta_2\right).$$

If at the end of the vacation there are three customers in the system, then the vacation belongs to the third level. At the beginning of service we have to come to the second level, this expectedly takes $\frac{\tau}{1-\varrho}$ time; then we are in the previous situation, so spending $\zeta_1+\zeta_2$ above the second level the busy period ends. The corresponding value is

$$f_3\left(\eta+\frac{\tau}{1-\varrho}-\tau+\zeta_1+\zeta_2\right).$$

If there are k customers present at the end of the vacation, the corresponding value is

$$f_k\left(\eta+(k-2)\frac{\tau}{1-\varrho}-\tau+\zeta_1+\zeta_2\right).$$

Consequently, the mean value of time spent above the second level for a busy period is

$$\begin{aligned}
\zeta_2' &= f_1\zeta_2+f_2\left(\zeta_1+\zeta_2\right)+\sum_{k=3}^{\infty} f_k\left(\eta+(k-2)\frac{\tau}{1-\varrho}-\tau+\zeta_1+\zeta_2\right)= \\
&= \zeta_2+\left(1-f_1\right)\zeta_1+\left(1-f_1-f_2\right)\left(\eta-\tau\right)+\frac{\tau}{1-\varrho}\sum_{k=3}^{\infty}\left(kf_k-2f_k\right)= \\
&= \zeta_2+\left(1-f_1\right)\zeta_1+\left(1-f_1-f_2\right)\left(\eta-\tau\right)+\frac{\tau}{1-\varrho}\left(\alpha+\alpha\lambda\eta-2+f_1\right).
\end{aligned}$$

The mean value of time spent on the second level

$$\xi_2' = \zeta_1' - \zeta_2' = f_2\left(\eta - \tau\right) - \frac{\tau}{c_0} - \zeta_2 - \left(1 - f_1\right) + \frac{\tau}{1-\varrho}\left(2 - f_1\right) =$$
$$= \xi_2 + \left(2 - f_1\right)\left(\frac{\tau}{1-\varrho} - \zeta_1\right) - \frac{\tau}{c_0} + f_2\left(\eta - \tau\right).$$

We are going to prove the recurrence relation by induction. Suppose that our reasoning is valid for the $i-1^{\text{th}}$ level. First we find the mean value of time spent above the i^{th} level. We have to consider several different cases similar to the previously examined ones.

If there is only one customer at the end of the vacation, the mean value is

$$f_1\zeta_i.$$

If there are two customers present at the end of the vacation, spending ζ_{i-1} above the i^{th} level we come to the first level and will be in the prevoious situation. The corresponding value is

$$f_2\left(\zeta_{i-1} + \zeta_i\right).$$

If there are i customers present at the end of the vacation, spending ζ_1 above the i^{th} level we come to the $i-1^{\text{th}}$ level then spending ζ_2 above the i^{th} level we come to level $i-2$, etc., and finally starting from the first level and spending ζ_i above the i^{th} one the busy period ends. The corresponding value is

$$f_i\left(\zeta_1 + \zeta_2 + \ldots + \zeta_{i-1} + \zeta_i\right).$$

If at the end of the vacation there are $i+1$ customers in the system, then the vacation belongs to the $i+1^{\text{th}}$ level. At the beginning of service we have to come to the i^{th} level, this expectedly takes $\frac{\tau}{1-\varrho}$ time; then we are in the previous situation, so spending $\zeta_1 + \zeta_2 + \ldots + \zeta_{i-1} + \zeta_i$ above the i^{th} level the busy period ends. The corresponding value is

$$f_{i+1}\left(\eta + \frac{\tau}{1-\varrho} - \tau + \zeta_1 + \zeta_2 + \ldots + \zeta_i\right).$$

If there are k customers at the end of the busy period ($k > i$), then reasoning the same way as above, the corresponding value is

$$f_k \left(\eta + (k-i) \frac{\tau}{1-\varrho} - \tau + \zeta_1 + \zeta_2 + \ldots + \zeta_i \right).$$

Adding these terms up, we get

$$\begin{aligned}
\zeta_i' = f_1 \zeta_i + f_2 (\zeta_{i-1} + \zeta_i) + \ldots + f_i (\zeta_1 + \zeta_2 + \ldots + \zeta_{i-1} + \zeta_i) + \\
+ \sum_{k=i+1}^{\infty} f_k \left(\eta + (k-i) \frac{\tau}{1-\varrho} - \tau + \zeta_1 + \zeta_2 + \ldots + \zeta_i \right) = \\
= \zeta_i + (1 - f_1) \zeta_{i-1} + \ldots + (1 - f_1 - \ldots - f_{i-1}) \zeta_1 + \\
+ (1 - f_1 - \ldots - f_i)(\eta - \tau) + \frac{\tau}{1-\varrho} \sum_{k=i+1}^{\infty} f_k (k-i) = \\
= \zeta_i + \sum_{j=1}^{i-1} (1 - f_1 - \ldots - f_j) \zeta_{i-j} + (1 - f_1 - \ldots - f_i)(\eta - \tau) + \\
+ \frac{\tau}{1-\varrho} (\alpha + \alpha\lambda\eta - i + (i-1) f_1 + (i-2) f_2 + \ldots + 2 f_{i-2} + f_{i-1})
\end{aligned}$$

In the case of the $i-1^{\text{th}}$ level we have

$$\begin{aligned}
\zeta_{i-1}' = \zeta_{i-1} + \sum_{j=1}^{i-2} (1 - f_1 - \ldots - f_j) \zeta_{i-1-j} + (1 - f_1 - \ldots - f_{i-1})(\eta - \tau) + \\
+ \frac{\tau}{1-\varrho} (\alpha + \alpha\lambda\eta - (i-1) + (i-2) f_1 + (i-3) f_2 + \ldots + 2 f_{i-3} + f_{i-2}).
\end{aligned}$$

The mean value of time spent on the i^{th} level is

$$\begin{aligned}
\xi_i' = \zeta_{i-1}' - \zeta_i' = \zeta_{i-1} - \zeta_i + \sum_{j=1}^{i-2} (1 - f_1 - \ldots - f_j)(\zeta_{i-1-j} - \zeta_{i-j}) - \\
- (1 - f_1 - \ldots - f_{i-1}) \zeta_1 + f_i (\eta - \tau) + \frac{\tau}{1-\varrho} (1 - f_1 - f_2 - \ldots - f_{i-1}) =
\end{aligned}$$

$$= \xi_i + \sum_{j=1}^{i-2} (1 - f_1 - \ldots - f_j)\, \xi_{i-j} + (1 - f_1 - \ldots - f_{i-1}) \left(\frac{\tau}{1-\varrho} - \zeta_1 \right) +$$
$$+ f_i\, (\eta - \tau) = \xi_i + \sum_{k=2}^{i-1} (1 - f_1 - \ldots - f_{i-k})\, \xi_k +$$
$$+ (1 - f_1 - \ldots - f_{i-1}) \left(\frac{\tau}{1-\varrho} - \zeta_1 \right) + f_i\, (\eta - \tau)\,,$$

which proves the theorem. ■

Bibliography

[1] ALMÁSI B., BOLCH G., SZTRIK J.: Heterogeneous finite-source retrial queues. *Journal of Mathematical Sciences*, **121** (2004) 2590–2596

[2] ALMÁSI B., SZTRIK J.: Reliability investigation of heterogeneous terminal systems using MOSEL. *Journal of Mathematical Sciences*, **123** (2004) 3795–3801

[3] ALMÁSI B., ROSZIK J., SZTRIK J.: Homogeneous finite-source retrial queues with server subject to breakdowns and repairs. *Mathematical and Computer Modelling*, **42** (2005) 673–682

[4] BRIÈRE G., CHAUDHRY M.L.: Computational analysis of single-server bulk-arrival queues $M^X/G/1$. *Comput. Oper. Res.*, Vol. 15, Nr. 3. (1988) 283-292

[5] BRIÈRE G., CHAUDHRY M.L.: Computational analysis of single-server bulk-service queues $M/G^Y/1$. *Adv. Appl. Prob.* **21** (1988) 207-225

[6] CHAUDHRY M.L., GUPTA U.C., AGARWAL M.: Exact and approximate solutions to steady-state single-server queues: $M/G/1$ – A unified approach. *Queueing Systems* **10** (1992) 351–380

[7] FALIN G.I., TEMPLETON J.G.C.: Retrial queues. Monographs on Statistics and Applied Probability **75** *Chapman & Hall* (1997)

[8] FARKAS G.: Investigation of a continuous cyclic-waiting problem by simulation. *Annales Univ. Sci. Budapest., Sect. Comp.* **19** (2000) 225–235

[9] Farkas G.: Numerical investigation of a cyclic-waiting queueing system with two types of customers. *Annales Univ. Sci. Budapest., Sect. Comp.* **21** (2002) 153–163

[10] Farkas G., Abdalla W.S.: Numerical investigation of the convergence to the limit distribution in a cyclic-waiting system. *Annales Univ. Sci. Budapest., Sect. Comp.* **20** (2001) 207–220

[11] Farkas G., Kárász P.: Investigation of a discrete cyclic-waiting problem by simulation. *Acta Acad. Paed. Agriensis, Sectio Mathematicae* **27** (2000) 57–62

[12] Gaver D.P., Jacobs P.A.: Time-limited tasks with uncertain outcomes. *XXI Seminar on Stability Problems of Stochastic Models, Eger, Hungary* (2001) 75–76

[13] Гихман И.И., Скороход А.В.: Введение в теорию случайных процессов. *изд. «Наука», Москва* (1977)

[14] Gnedenko B.V., Kovalenko I.N.: Introduction to queuing theory. *Birkhäuser, Boston* (1989)

[15] Гнеденко Б.В., Коваленко И.Н.: Введение в теорию массового обслуживания. *изд. «КомКнига», Москва* (2005)

[16] Györfi L., Páli I.: Tömegkiszolgálás informatikai rendszerekben. *Műegyetemi Kiadó, Budapest* (2001)

[17] Heyman D.P., Sobel M.J.: Stochastic models. Handbooks in Operations Research and Management Science, II. *North-Holland* (1990)

[18] Ивченко Г.И., Каштанов В.А., Коваленко И.Н.: Теория массового обслуживания. *изд. «Высшая школа», Москва* (1982)

[19] Kárász P.: On a class of retrial systems with uniformly distributed service time. *Magdeburger Stochastic Tage, Magdeburg, Germany*, Otto von Guericke Universität, Magdeburg (2002) 110–112

[20] Kárász P.: On a class of cyclic-waiting queuing systems with refusals. *Conference of PhD Students in Computer Science, Szeged, Hungary*, Szeged University (2002) 55–56

[21] Kárász P.: Special retrial systems with requests of two types. *International Summer Seminar "Stochastic Dynamical Systems", Sudak, Crimea, Ukraine*, Institute of Mathematics of the National Academy of Sciences of Ukraine, Kiev (2003) 29–30

[22] Kárász P.: Special retrial systems with exponentially and uniformly distributed service times. *International Conference in Memoriam John von Neumann, Budapest, Hungary*, John von Neumann Computer Society, Budapest Polytechnic, Hungarian Fuzzy Association, Budapest (2003) 199–207

[23] Kárász P.: Special retrial systems with requests of two types. *Theory of Stochastic Processes* **10 (26)** 3-4 (2004) 51–56

[24] Kárász P.: $M/G/1$ Queuing System with Two Types of Vacation. *Conference of PhD Students in Computer Science, Szeged, Hungary*, Szeged University (2004) 66

[25] Kárász P.: Equilibrium distribution for bulk-arrival $M/G/1$ system with vacation. *XXIV International Seminar on Stability Problems for Stochastic Models, Jurmala, Latvia*, Transport and Telecommunications Institute, Riga (2004) 133

[26] Kárász P.: Ciklikus várakozási idejű tömegkiszolgálási rendszerek diszkrét eloszlások esetén. *XXVII. Magyar Operációkutatási Konferencia, Balatonöszöd, Hungary*, Magyar Operációkutatási Társaság, Budapest (2007) 45

[27] Kárász P.: Investigating the behaviour of a discrete retrial system. *Acta Cybernetica* **18** 4 (2008)

[28] Kárász P.: Equilibrium distribution for bulk-arrival $M/G/1$ system with vacation. *(submitted to Journal of Mathematical Sciences)*

[29] Kárász P.: A special discrete cyclic-waiting queuing system. *(to be published in Central European Journal of Operations Research)*

[30] Kárász P., Farkas G.: Exact solution for a two-type customers retrial system. *Computers and Mathematics with Applications* **49** (2005) 95–102

[31] Kárász P., Farkas G.: Simulation of Lakatos-type queuing systems. *5th St.Petersburg Workshop on Simulation, St.Petersburg, Russia*, St.Petersburg State University (2005) 333–335

[32] Kendall D.G.: Some problems in the theory of queues. *Journal of the Royal Statistical Society, Series B* **13** (1951) 151–185

[33] Kleinrock L.: Theory. Queueing Systems I. *John Wiley & Sons* (1975)

[34] Коба Е.В.: О системе обслуживания $GI/G/1$ с повторением заявок при обслуживании в порядке очереди. *Доповіді НАН України* № **6** (2000) 101–103

[35] Коба Е.В.: Об условии устойчивости системы обслуживания $M/D/1$ с повторяющимися заявками и ограниченным временем ожидания. *Кибернетика и системный анализ* № **2** (2000) 184–187

[36] Koba E.V.: Stability condition for $M/D/1$ retrial queuing system with a limited waiting time. *Cybernetics and Systems Analysis* **36** 2 (2000) 313–315

[37] Коба Е.В.: Система обслуживания $M/D/1$ с заявками, повторяющимися через постоянное время, при частичной синхронизации входящего потока. *Кибернетика и системный анализ* № **6** (2000) 177–180

[38] Koba E.V.: An $M/D/1$ Queueing system with partial synchronization of its incoming flow and demands repeating at constant intervals. *Cybernetics and Systems Analysis* **36** 6 (2000) 946–948

[39] Коба Е.В.: Об одной системе обслуживания с повторением и выталкиванием заявок. *Проблемы управления и информатики* № **1** (2001) 58–62

[40] Коба E.V.: Stability and ergodicity conditions for a $GI/G/1$ retrial queueing system with FIFO queueing discipline. *Int. Gnedenko Conference, Kiev, Ukraine* (2002) 51

[41] Коба E.V.: On a $GI/G/1$ retrial queueing system with a FIFO queueing discipline. *Theory of Stochastic Processes* **8 (24)** 1-2 (2002) 201–207

[42] Коба Е.В.: Достаточное условие эргодичности системы $M/D/1$ с T-возвращением и приоритетом задержанных заявок. *Доповіді НАН України* № **5** (2003) 17–20

[43] Коба О.В.: Стаціонарні характеристики системи масового обслуговування $GI/G/1$ із T-поверненням при обслуговуванні в порядку черги. *Вiснук НАУ* № **1** (2003) 122–125

[44] Коба О.В.: Дослiдження систем обслуговування з поверненням заявок при неекспоненцiальному розподiлi часу перебування на орбiтi. Роздiл 3: Системи Ласло Лакатоша та їх узагальнення. D.Sc. thesis, *Институт Кибернетики, Киев* (2005) 116–158

[45] Коба О.В., Михалевич К.В.: Порiвняння систем типу $M/M/1$ з швидким поверненням заявок при рiзних дисциплинах обслуговування. *System Research & Information Technologies* **2** (2003) 59–68

[46] Коба О.В., Михалевич К.В.: Сравнение систем обслуживания типа $M/G/1$ с повторением при быстром возвращении с орбиты. *Modern Mathematical Methods of Telecommunication Networks, Gomel, Byelorussia*, BGU, Minsk (2003) 136–138

[47] Коваленко И.Н.: Некоторые вопросы теории надежности сложных систем. Кибернетику – на службу коммунизму **2**, *изд. Энергия, Москва* (1964) 194–205

[48] Коваленко И.Н.: Некоторые аналические методы в теории массового обслуживания. Кибернетику – на службу коммунизму **2**, *изд. Энергия, Москва* (1964) 325–337

[49] Коваленко И.Н.: Вероятность потери в системе обслуживания $M/G/m$ с T-повторением вызовов в режиме малой нагрузки. *Доповiдi НАН України* № **5** (2002) 77–80

[50] Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф.: Справочник по теории вероятностей и математической статистике. *изд. «Наукова думка», Киев* (1978)

[51] Laevens K., Bruneel H.: Analysis of a single-wavelength optical buffer. *Infocom 2003, San Francisco, United States* (2003) 105–113

[52] Laevens K., Van Houdt B., Blondia C., Bruneel H.: On the sustainable load of fiber delay line buffers. *Electronic Letters* **40** (2004) 137–138

[53] Lakatos L.: On a simple continuous cyclic-waiting problem. *Annales Univ. Sci. Budapest., Sect. Comp.* **14** (1994) 105–113

[54] Lakatos L.: On a cyclic-waiting queuing system. *Theory of Stochastic Processes* **2 (18)** 1-2 (1996) 176–180

[55] Lakatos L.: Ergodic distribution for the $M/G/1$ system with vacation. *Theory of Stochastic Processes* **3 (19)** 1-2 (1997) 244–255

[56] Lakatos L.: On a simple discrete cyclic-waiting queuing problem. *J. Math. Sci. (New York)*, 92 (4) (1998) 4031–4034

[57] Lakatos L.: Equilibrium distributions for the $M/G/1$ and related systems. *Publicationes Mathematicae* Vol. 55, Nr. 1-2. (1999) 123-140.

[58] Lakatos L.: On the $M^X/G/1$ system. *Annales Univ. Sci. Budapest., Sect. Comp.* **18** (1999) 137–150

[59] Lakatos L.: On the $M/G/1$ system with bulk arrivals and vacation. *Theory of Stochastic Processes* **5 (21)** 3-4 (1999) 127–136

[60] Lakatos L.: Limit distributions for some cyclic-waiting queuing systems. *Ukrainian Math. Congress 2001 Proc., Sect. 9. Probability theory and math. stat., Institute of Mathematics, Kiev* (2002) 102–106

[61] Lakatos L.: A special cyclic-waiting queuing system with refusals. *J. Math. Sci. (New York)*, 111 (3) (2002) 3541–3544

[62] Lakatos L.: A retrial system with time-limited tasks. *Theory of Stochastic Processes* **8 (24)** 3-4 (2002) 250–256

[63] Lakatos L.: A special cyclic-waiting queueing system with refusals: The discrete time case. *Annales Univ. Sci. Budapest. Sect. Comp.* **24** (2004) 323–333

[64] Лакатош Л.: Система $M/G/1$ с настройкой в начале периода занятости. *Кибернетика и системный анализ* № **1** (2005) 117–123

[65] Lakatos L.: A retrial queueing system with urgent customers. *J. Math. Sci.* **138** 1 (2006) 5405–5409

[66] Lakatos L.: A note on the Pollaczek–Khinchin formula. *Annales Univ. Sci. Budapest. Sect. Comp.* **29** (2008) 83–91

[67] Lakatos L., Koltai T.: A discrete retrial system with uniformly distributed service time. *Annales Univ. Sci. Budapest., Sect. Comp.* **22** (2003) 225–234

[68] Lakatos L., Szeidl L., Telek M.: Tömegkiszolgálás. In: Informatikai algoritmusok 2. *ELTE Eötvös Kiadó, Budapest* (2005) 1298–1346

[69] Lakatos L., Zbăganu G.: Waiting time in cyclic-waiting systems. *Annales Univ. Sci. Budapest., Sect. Comp.* **27** (2007) 217–228

[70] Mykhalevych K.V.: A comparison of a classical retrial $M/G/1$ queueing system and a Lakatos-type $M/G/1$ cyclic-waiting time queueing system. *Annales Univ. Sci. Budapest., Sect. Comp.* **23** (2004) 229–238

[71] Прохоров Ю.В., Розанов Ю.А.: Теория вероятностей. Основные понятия, предельные теоремы, случайные процессы. *изд. «Наука», Москва* (1973)

[72] Rogiest W., Laevens K, Walraevens J., Bruneel H.: Analyzing a degenerate buffer with general inter-arrival and service times in discrete time. *Queueing Systems* **56** (2007) 203–212

[73] Roszik J., Sztrik J., Kim C-S.: Retrial queues in the performance modeling of cellular mobile networks using MOSEL. *International Journal of Simulation: Systems, Science & Technology* **6** (2005) 38–47

[74] Sztrik J.: Tool supported performance modelling of finite-source retrial queues with breakdowns. *Publicationes Mathematicae*, **66** (2005) 197–211

[75] Sztrik J., Almási B., Roszik J.: Heterogeneous finite-source retrial queues with server subject to breakdowns and repairs. *Journal of Mathematical Sciences*, **132** (2006) 677–685

[76] Takagi H.: Queueing analysis: a foundation of performance analysis. *North-Holland* (1991)

[77] Tijms H.C.: Stochastic models. An algorithmic approach. *John Wiley & Sons, Chichester* (1994)

Summary

Cyclic-waiting queuing systems are charactarized by the feature that a request for service can be repeated only after a constant period of time. A typical example can be an automatic redialling device, which again and again attempts to dial the called number after a deterministic time if the line is engaged. This also occurs at airports where planes can start landing upon arrival or have to start a circular manoeuvre, when the runway is used or there are other planes waiting, and they can only put their next request to land after they have completed a full cycle. Another application in digital technology is the use of an optical buffer, which is a device that is capable of temporarily storing light (or rather, data in the form of light). As light cannot be frozen, a typical optical buffer is realized by a single loop, in which data circulate a variable number of times.

The dissertation describes and further generalizes the results of Lakatos on cyclic-waiting systems. It investigates the so-called relative priority systems, which serves two types of customers of different priority. In these systems only one customer of first type can be present, it can only be accepted for service in the case of a free system, whereas in all other cases the requests of such customers are turned down. There is no such restriction on customers of second type; they are serviced immediately or join a queue in case of a busy server. This model can be applied for systems accepting impatient customers who need urgent service; if they cannot get serviced, they leave the system and find another server which is free.

Both continuous and discrete systems are considered. In the first case inter-arrival times are exponentially, service times are either exponentially or uniformly distributed; whereas in the latter one inter-arrival times are

geometrically, service times are either geometrically or uniformly distributed random variables. For discrete systems three collision disciplines are considered. In both cases an embedded Markov-chain is defined, and the generating functions of transition probabilities, as well as that of the equilibrium probabilities are given. Necessary and sufficient conditions of ergodicity are provided for each case. To validate results, equilibrium probabilities of states 0 and 1 are taken a closer look, and results of numerical computer experiments are included. The expected value of the queue-length is also given, it is explicitly determined in the case of geometric service time distributions, and its dependence on input parameters is analyzed graphically. In both chapters the unified theory is given, i. e. generating functions can be combined together as necessary, in accordance with the type of service time distribution of first- and second-type customers. Similarly, the generating function of the equilibrium distribution, as well as the conditions of ergodicity are valid for all cases.

A queuing system with vacation is a queuing system in which the server intermittently spends time away from the queue, perhaps because of a breakdown and repair or because it is serving other jobs. Some examples are token-passing schemes in local area networks with distributed channel access control, and single-server multi-queue models.

The last chapter of the dissertation is devoted to the investigation of a queuing system which accepts bulk-arrivals, and where the first appearing group of requests initiates a vacation during which the server is prepared for the service, and actual service can only begin after this period of time. The dissertation presents recursive formulae for the equilibrium distribution of this system.

Összefoglalás

A ciklikus várakozási idejű tömegkiszolgálási rendszerek jellemzője, hogy a beérkező egységek kiszolgálásukat csak bizonyos konstans idő, illetve ennek többszörösei eltelte után igényelhetik. Példaként egy automatikus újratárcsázó készülék említhető, amely újból és újból tárcsázza a hívott számot egy meghatározott idő elteltével, amíg az foglalt. Hasonló mechanizmussal találkozhatunk repülőtereken, ahol a repülőgépek érkezésükkor vagy leszállnak, vagy egy körpályára kényszerülnek, ha a kifutópálya foglalt, és legközelebb egy teljes kör megtétele után kezdhetik meg a leszállást. Egy digitális technológiai alkalmazás pedig az optikai pufferek használata, amely a fény (helyesebben fény formájában kódolt adatok) tárolására képes. Mivel a fényt megállítani nem lehet, az optikai puffer tipikus modellje egy egyszerű hurok, amelyben a fény keringhet.

Az értekezés Lakatos fenti típusú rendszerekre vonatkozó eredményeit tárgyalja és ebből kiindulva általánosítja azokat. Egy úgynevezett relatív prioritásos rendszert vizsgálunk, amely két, különböző prioritású igényt szolgál ki. Ezen rendszerekben az egyes típusú igényből csak egy lehet jelen, csak szabad kiszolgáló egység esetén fogadjuk be, minden más esetben elutasítjuk őket. A kettes típusú igényekre hasonló megszorítás nem érvényes, őket azonnal kiszolgáljuk vagy a várakozási sor végére kerülnek. A modell olyan „türelmetlen" igényeket is fogadó rendszerekre alkalmazható, amelyeknek sürgős kiszolgálása szükségeltetik; ha a kiszolgálás nem lehetséges, elhagyják a rendszert annak reményében, hogy egy másik, szabad kiszolgáló egységet találnak.

A tárgyalás során mind folytonos, mind diszkrét rendszereket vizsgálunk. Az első esetben a beérkezési időköz exponenciális, a kiszolgálási idő exponenciális vagy egyenletes eloszlású; míg az utóbbi esetben a beérkezési időköz geometriai, a kiszolgálási idő pedig geometriai vagy egyenletes eloszlású valószínűségi változó. Diszkrét rendszereknél háromféle ütközéskezelési módszert tekintünk. Mindkét fejezetben egy beágyazott Markov-láncot definiálunk, az átmenet-valószínűségek és az egyensúlyi állapot valószínűségeinek generátorfüggvényeit is megadjuk. Az ergodicitás szükséges és elégséges feltételeit minden esetben meghatározzuk. Az eredmények hitelesítésére az egyensúlyi eloszlás 0 és 1 állapotainak valószínűségeit külön is elemezzük, illetve számítógéppel végzett numerikus eredményeket is közlünk. A sorhossz várható értékét meghatározzuk, geometriai eloszlású kiszolgálási idő esetén ezt expliciten megadjuk, a bemenő paraméterektől való függését pedig grafikusan is ábrázoljuk. Mindkét fejezetben az egyesített elmélet eredményeit adjuk közre, azaz a feltüntetett generátorfüggvények tetszőlegesen kombinálhatók az egyes és kettes típusú kiszolgálási idő eloszlásoknak megfelelően. Hasonlóképpen, az ergodicitási feltételek is bármely esetre érvényesek.

A vakációs tömegkiszolgálási rendszerek olyan sorbanállási rendszerek, amelyekben a kiszolgáló egység időről időre szünetelteti a sor kiszolgálását, meghibásodás, javítás vagy más feladat végrehajtása miatt. Ilyen típussal többek között LAN hálózatokban, az osztott csatornahozzáférésű „token-passing" topológia, vagy az egykiszolgálós, többsoros rendszerek esetén találkozhatunk.

Az értekezés utolsó fejezetét olyan vakációs rendszerek vizsgálatának szenteljük, amelyek csoportos beérkezéseket fogadnak. Ezekben az első beérkező csoport egy vakációs periódust indít, ami alatt a kiszolgáló egységet előkészítik a működésre, és a tényleges kiszolgálás csak ez után kezdődhet el. Meghatározzuk ezen rendszer egyensúlyi eloszlását rekurzív formulák segítségével.

Acknowledgements

In the first place I would like to thank my supervisor, László Lakatos his invaluable and generous help and attentive guidance all through my studies and my way to the complete dissertation.

Thanks should go to my workplace, Budapest Tech for their financial support to my studies; to Prof. Dezső Sima former, and Prof. László Szeidl present deans of John von Neumann Faculty of Informatics for their professional help; and to my colleagues for their encouragement, beneficial remarks and for covering my classes while I was away on conferences.

I owe a lot to my loving family and friends for gently pressing on me, for their patience and for backing me even in the most hopeless moments.

And at last but not least I shall always be grateful to my school teacher, Mária Csiszár, who led my extracurricular classes at primary school, then she taught me at grammar school, and who thus has been guiding me through the maze of mathematics for 7 years, for the longest period of time in my life. I have learnt the precise mathematical way of thinking and logical arguing that conducts me upto present days from her.

Köszönetnyilvánítás

Az első helyen szeretném megköszönni témavezetőmnek, Lakatos Lászlónak az értekezés elkészítéséhez nyújtott felbecsülhetetlen és önzetlen segítségét valamint mindenre kiterjedő, alapos irányítását a tanulmányaim során.

Köszönet illeti munkahelyemet, a Budapesti Műszaki Főiskolát a tanulmányaimhoz nyújtott anyagi támogatásért; a Nemann János Informatikai Kar korábbi és jelenlegi dékánját, Prof. Sima Dezsőt és Prof. Szeidl Lászlót a szakmai segítségért; valamint kollégáimat a bátorításért, a hasznos megjegyzésekért és az óráim helyettesítéséért amíg konferenciákon vettem részt.

Sokkal tartozom szerető családomnak és barátaimnak a szelíd noszogatásért, a türelemért és a támaszért még a legreménytelenebb pillanatokban is.

Végül, de nem utolsó sorban tisztelettel és hálával tartozom egykori matematika tanárnőmnek, Csiszár Máriának, akinek kezdetben az általános iskolai matematika szakköreit látogattam, majd a gimnáziumban is tanárom volt, és így 7 évig, életem során a leghosszabb ideig kalauzolt a matematika útvesztőjében. Tőle tanultam azt a matematikai precizitást és logikus gondolkodásmódot, amely a mai napig vezérel.